

Meningkatkan Literasi Penilaian Kelas

Prinsip-prinsip dan Langkah-langkah Praktis

Buku ini diperuntukkan kepada guru dan calon guru secara umum dan khususnya guru dan calon guru mata pelajaran Bahasa. Buku ini ditulis berdasarkan hasil penelitian yang dilakukan penulis serta referensi dan praktik-praktik penilaian terbaik dari dalam dan luar negeri.

Buku ini memuat 10 bab yang berkaitan dan dimulai pemahaman teori penilaian, prinsip-prinsip penilaian dan pengajaran, kasus dalam penilaian, teknik dalam penilaian, apa saja yang harus dinilai, hingga strategi penilaian yang dapat diterapkan dalam penilaian Bahasa. Pada setiap bab dimuat latihan soal yang berguna untuk mengetahui pemahaman pembaca.

Buku ini merupakan salah satu luaran Program Hibah Penelitian Kementristekdikti tahun 2019. Penulis sangat menghargai kesempatan yang telah diberikan kepada penulis melalui program ini.


Rajawali Pers
PT RAJAWALI PERS
Jl. Raya Lawangrangin No. 112
Kel. Lawangrangin, Kec. Tembung, Kota Depok, 16156
Telp: 021-8431113, Fax: 021-8431116
Email: rajan@rajawali.pers.id
www.rajawali.pers.id

RAJAWALI PERS
DIVISI BUKU PERGURUAN TINGGI
PENDIDIKAN



Meningkatkan **Literasi Penilaian Kelas** Prinsip-prinsip dan Langkah-langkah Praktis Penyusun

Siti Zulaiha | Herri Mulyono



Meningkatkan Literasi Penilaian Kelas

Prinsip-prinsip dan Langkah-langkah Praktis



Siti Zulaiha | Herri Mulyono

Meningkatkan
Literasi
Penilaian Kelas
Prinsip-prinsip dan Langkah-langkah Praktis

Meningkatkan
Literasi
Penilaian Kelas
Prinsip-prinsip dan Langkah-langkah Praktis

Siti Zulaiha | Herri Mulyono



RAJAWALI PERS
Divisi Buku Perguruan Tinggi
PT RajaGrafindo Persada
D E P O K

Perpustakaan Nasional: Katalog dalam terbitan (KDT)

Siti Zulaiha dan Herri Mulyono

Meningkatkan Literasi Penilaian Kelas: Prinsip-prinsip dan Langkah-langkah Praktis Penyusun/Siti Zulaiha dan Herri Mulyono
—Ed. 1, Cet. 1.—Depok: Rajawali Pers, 2020.

x, 90 hlm., 23 cm.

Bibliografi: 87

ISBN 978-623-231-436-8

Hak cipta 2020, pada penulis

Dilarang mengutip sebagian atau seluruh isi buku ini dengan cara apa pun, termasuk dengan cara penggunaan mesin fotokopi, tanpa izin sah dari penerbit

2020.2662 RAJ

Siti Zulaiha

Herri Mulyono

***MENINGKATKAN LITERASI PENILAIAN KELAS:
Prinsip-prinsip dan Langkah-langkah Praktis***

Cetakan ke-1, Juni 2020

Hak penerbitan pada PT RajaGrafindo Persada, Depok

Editor : Prajna Vita

Setter : Feni Erviana

Desain Cover : Tim Kreatif RGP

Dicetak di Rajawali Printing

PT RAJAGRAFINDO PERSADA

Anggota IKAPI

Kantor Pusat:

Jl. Raya Leuwinanggung, No.112, Kel. Leuwinanggung, Kec. Tapos, Kota Depok 16956

Telepon : (021) 84311162

E-mail : rajapers@rajagrafindo.co.id

<http://www.rajagrafindo.co.id>

Perwakilan:

Jakarta-16956 Jl. Raya Leuwinanggung No. 112, Kel. Leuwinanggung, Kec. Tapos, Depok, Telp. (021) 84311162. **Bandung**-40243, Jl. H. Kurdi Timur No. 8 Komplek Kurdi, Telp. 022-5206202. **Yogyakarta**-Perum. Pondok Soragan Indah Blok A1, Jl. Soragan, Ngestiharjo, Kasihan, Bantul, Telp. 0274-625093. **Surabaya**-60118, Jl. Rungkut Harapan Blok A No. 09, Telp. 031-8700819. **Palembang**-30137, Jl. Macan Kumbang III No. 10/4459 RT 78 Kel. Demang Lebar Daun, Telp. 0711-445062. **Pekanbaru**-28294, Perum De' Diandra Land Blok C 1 No. 1, Jl. Kartama Marpoyan Damai, Telp. 0761-65807. **Medan**-20144, Jl. Eka Rasmi Gg. Eka Rossa No. 3A Blok A Komplek Johor Residence Kec. Medan Johor, Telp. 061-7871546. **Makassar**-90221, Jl. Sultan Alauddin Komp. Bumi Permata Hijau Bumi 14 Blok A14 No. 3, Telp. 0411-861618. **Banjarmasin**-70114, Jl. Bali No. 31 Rt 05, Telp. 0511-3352060. **Bali**, Jl. Imam Bonjol Gg.100/V No. 2, Denpasar Hp. 081222805496. **Bandar Lampung**-35115, Perum. Bilabong Jaya Block B8 No. 3 Susunan Baru, Langkapura, Hp. 081299047094.

A decorative header featuring a large pencil pointing downwards on the left. The background is filled with various educational icons in shades of gray, including gears, books, a microscope, a clock, a graduation cap, a DNA helix, a ruler, a paper airplane, a paperclip, a heart with a cross, a building, and the letters 'A', 'B', and 'C'.

Prakata

Puji dan syukur kehadiran Allah Swt., yang telah melimpahkan rahmat dan karuniaNya sehingga kami telah berhasil menyelesaikan proses penulisan buku ini. Buku yang berjudul “*Meningkatkan Literasi Penilaian Kelas: Prinsip-prinsip dan Langkah-langkah Praktis*” bertujuan untuk menambah referensi praktis dalam penilaian kelas. Buku ini diperuntukkan kepada guru dan calon guru secara umum dan khususnya guru dan calon guru mata pelajaran Bahasa. Buku ini ditulis berdasarkan hasil penelitian yang dilakukan penulis serta referensi dan praktik-praktik penilaian terbaik dari dalam dan luar negeri.

Buku ini memuat 10 bab yang berkaitan dan dimulai dengan pemahaman teori penilaian, prinsip-prinsip penilaian dan pengajaran, kasus dalam penilaian, teknik dalam penilaian, apa saja yang harus dinilai, hingga strategi penilaian yang dapat diterapkan dalam penilaian Bahasa. Pada setiap bab dimuat latihan soal yang berguna untuk mengetahui pemahaman pembaca.

Buku ini merupakan salah satu luaran Program Hibah Penelitian Kemenristekdikti tahun 2019. Penulis sangat menghargai kesempatan yang telah diberikan kepada penulis melalui program ini.

Kami mengucapkan terima kasih kepada seluruh pihak yang telah mendukung dalam penyusunan buku ini. Harapan kami buku ini bermanfaat untuk meningkatkan kemampuan guru dan calon guru

dalam bidang penilaian kelas sehingga meningkat pula mutu pencapaian siswa. Kami terbuka dengan kritik dan saran dari pembaca.

Jakarta, 17 September 2019

Penulis



Daftar Isi

KATA PENGANTAR	v
DAFTAR ISI	vii
BAB 1 PENDAHULUAN	1
A. Definisi Literasi Penilaian	1
B. Literasi Penilaian dan Pengembangan Keprofesian Guru	3
BAB 2 PRINSIP-PRINSIP PENILAIAN	5
A. Praktis	6
B. Andal/Reliabel	7
C. Sahih/Valid	9
D. Autentik	13
E. <i>Washback</i>	13
Latihan	14
BAB 3 PENILAIAN DAN PENGAJARAN	15
A. Penilaian Formal dan Informal	16
B. Penilaian Formatif dan Sumatif	17

C. <i>Norm-Referenced</i> dan <i>Criterion-Referenced Testing</i>	18
Latihan	19
BAB 4 PENILAIAN KELAS	21
A. Penilaian Pembelajaran (<i>Assessment of Learning</i>)	22
B. Penilaian untuk Pembelajaran (<i>Assessment for Learning</i>)	23
C. Penilaian Sebagai Pembelajaran (<i>Assessment as Learning</i>)	26
Latihan	28
BAB 5 MERANCANG TES	29
A. Siklus Perancangan Tes	30
B. Definisi Konstruk Model	31
C. Model Kompetensi Komunikatif	32
D. Mendesain Spesifikasi Tes	33
E. Pelaksanaan	35
Latihan	35
BAB 6 BENTUK-BENTUK PENILAIAN	37
A. Penilaian Respons Terpilih (<i>Selected Response</i>)	38
B. Penilaian Respons Terstruktur (<i>Constructed Response</i>)	40
C. Penilaian Respons Pribadi (<i>Personal Response</i>)	43
Latihan	50
BAB 7 PENILAIAN ALTERNATIF	51
A. Apa itu Penilaian Alternatif?	52
B. Mengapa Kita Harus Menggunakan Penilaian Alternatif?	52
C. Kelebihan dan Kekurangan dari Penilaian Alternatif	53
D. Pedoman untuk Membangun Penilaian Alternatif	53
Latihan	54

BAB 8 MERANCANG DAN MENGGUNAKAN RUBRIK	55
A. Pengertian Rubrik	55
B. Manfaat Rubrik	56
C. Jenis-jenis Rubrik	58
D. Komponen-komponen Rubrik	62
E. Langkah-langkah Merancang Rubrik	64
F. Me-review Rubrik	65
G. Menggunakan Rubrik	66
Latihan	67
BAB 9 MENILAI KETERAMPILAN BERPIKIR TINGKAT TINGGI (HOTS)	69
A. Prinsip Umum dalam Menilai <i>Higher Order Thinking Skills</i> (HOTS)	71
B. Strategi dalam Memberikan <i>Feedback</i> dan Penskoran Tugas dalam HOTS	73
Latihan	77
BAB 10 UMPAN BALIK YANG EFEKTIF	79
A. Konsep Umpan Balik (<i>Feedback</i>)	79
B. Strategi dan Konten Umpan Balik	80
C. Teknik Mengajarkan Siswa Memanfaatkan Umpan Balik	83
Latihan	85
DAFTAR PUSTAKA	87
BODATA PENULIS	89

Bab 1

Pendahuluan

Objek Pembelajaran

- A. Definisi Literasi Penilaian
- B. Literasi Penilaian dan Pengembangan Keprofesian Guru

A. Definisi Literasi Penilaian

Telah diakui secara luas bahwa menilai kinerja siswa adalah salah satu elemen penting dari pekerjaan guru. Penilaian memang merupakan bagian integral dari proses belajar mengajar (Ashraf & Zolfaghari, 2018) dan bahwa guru menghabiskan banyak waktu untuk terlibat dalam kegiatan penilaian (DeLuca & Klinger, 2010). Penilaian pembelajaran bahasa dianggap tidak hanya sebagai sarana untuk memantau kemajuan siswa dan meningkatkan prestasi tetapi juga untuk mempromosikan pembelajaran (Earl, 2003). Dengan kata lain, penilaian mempunyai berbagai tujuan dan merupakan aspek penting dari praktik guru di kelas

dan kehidupan profesional mereka. Oleh karena itu, penting bagi guru untuk memiliki tingkat literasi penilaian yang memadai untuk dapat menilai pembelajaran siswa secara tepat.

Stiggins (1991) mendefinisikan literasi penilaian sebagai pemahaman prinsip-prinsip penilaian dan keterampilan yang dibutuhkan dalam menilai pembelajaran siswa. Stiggins (1995) lebih lanjut mencatat bahwa mereka yang melek penilaian tahu apa yang harus dinilai, bagaimana menilai, potensi masalah apa yang akan terjadi dan tahu apa yang harus dilakukan untuk mengatasi masalah tersebut. Literasi penilaian sejak itu menjadi istilah yang umum digunakan. Peneliti dan praktisi berusaha untuk mengonseptualisasikan makna literasi penilaian lebih lanjut. Sebagai contoh, (Djoub, 2017) menegaskan bahwa literasi penilaian berarti guru memiliki pengetahuan dan alat untuk mengetahui apa dan bagaimana menilai berdasarkan tujuan tertentu serta mengetahui keputusan apa yang harus diambil dalam menilai prestasi siswa. Dengan kata lain, penilaian guru mengacu pada pemahaman guru tentang prinsip-prinsip di balik memilih dan merancang tugas, menilai pekerjaan siswa, dan menafsirkan serta menggunakan data penilaian untuk mendukung pembelajaran siswa.

Konsep literasi penilaian yang dibahas di atas memperkuat pengaruh praktik penilaian guru terhadap kualitas pembelajaran siswa. Dikatakan bahwa guru yang mempunyai literasi penilaian yang baik mampu mengintegrasikan penilaian dengan pengajaran. Oleh karena itu, mereka mampu menarik kesimpulan tentang prestasi siswa secara akurat dan efisien, serta mampu mengomunikasikan hasil penilaian kepada pemangku kepentingan terkait (Brookhart, 1999). Singkatnya, guru yang melek penilaian mengetahui metode yang tepat untuk digunakan dalam mengumpulkan data yang dapat dipercaya tentang kinerja siswa, tahu cara menggunakan penilaian untuk mendukung pembelajaran siswa dan bagaimana mengomunikasikan hasil penilaian secara efektif dan akurat.

Namun pada kenyataannya proses asesmen tidak semudah yang dibayangkan oleh sebagian guru. Banyak guru yang merasa belum paham tentang prinsip-prinsip dari penilaian itu sendiri. Itulah mengapa literasi penilaian penting untuk dikuasai oleh guru-guru agar mereka paham dan mampu menerapkannya sesuai dengan prinsip-prinsip yang ada. Dalam beberapa tahun terakhir penelitian tentang literasi penilaian

telah dilakukan oleh beberapa ahli untuk mengetahui pengetahuan guru, keterampilan, dan pemahaman tentang dasar-dasar penilaian beserta praktiknya.

Beberapa penelitian dalam konteks yang berbeda menunjukkan berbagai temuan signifikan dari literasi penilaian guru. Di bidang pendidikan umum, misalnya, Yamtim & Wongwanich (2014) mengeksplorasi tingkat melek penilaian guru sekolah dasar di Thailand. Menggunakan Metlers (2003) Kuesioner Literasi Penilaian Kelas, terungkap bahwa guru memiliki tingkat literasi penilaian yang rendah. Dalam penelitian yang lebih baru, Rahman (2018) meneliti persepsi guru sains sekunder dan praktik penilaian kelas. Hasilnya menyarankan bahwa guru menganggap penilaian kelas sebagai penilaian pembelajaran dan bahwa apa yang mereka klaim untuk praktik tidak dipraktikkan di dalam kelas. Khusus untuk konteks pembelajaran bahasa asing, Jannati (2015) dalam penelitiannya menemukan bahwa guru menyadari prinsip-prinsip dasar penilaian. Namun, praktik guru tidak mencerminkannya. (Djoub, 2017) meneliti pengaruh literasi penilaian guru terhadap praktik guru. Data dikumpulkan dari para guru di seluruh dunia melalui survei online. Temuan menunjukkan bahwa guru kurang melek penilaian dan itu tercermin dalam praktik mereka.

B. Literasi Penilaian dan Pengembangan Keprofesian Guru

Melihat pentingnya literasi asesmen guru, diskusi-diskusi dan penelitian semakin berkembang dan berfokus pada pentingnya meningkatkan literasi asesmen guru. Beberapa pakar mengajukan beragam strategi untuk meningkatkan literasi guru. Coombe, Al-Hamly, & Troudi (2009) misalnya, menyatakan bahwa literasi asesmen dapat dilakukan melalui pemahaman apa itu asesmen yang baik, pengembangan keprofesian baik melalui moda daring maupun tatap muka pada lokakarya di semua level, komitmen untuk melakukan perubahan, serta penyediaan sumber belajar terutama online untuk mendukung pengembangan keprofesian. Strategi peningkatan literasi asesmen melalui pengembangan keprofesian juga disuarakan oleh Koh (2011). Hasil eksperimen Koh menunjukkan bahwa ada peningkatan literasi asesmen dari guru yang melakukan pengembangan keprofesian berkelanjutan. Dalam konteks yang berbeda, Odo (2015) menerapkan strategi tutorial individu terhadap calon guru di mana setiap guru mempunyai kesempatan untuk melakukan praktik

asesmen di bawah bimbingan instruktur dan melakukan refleksi pribadi terhadap praktik yang sudah dan akan dilakukan.

Paparan di atas mengemukakan pentingnya membekali guru dengan pengembangan keprofesian dan memberikan cukup waktu bagi guru untuk menerapkan apa yang mereka pelajari tentang asesmen. Oleh karena itu, lokakarya atau pelatihan singkat tentang asesmen tidaklah cukup. Kesuksesan sebuah pengembangan keprofesian untuk membangun literasi asesmen membutuhkan perubahan pada praktik pendidikan kita serta komitmen dari guru dan seluruh pemangku kepentingan (Koh, 2011). Lebih lanjut lagi, pengembangan keprofesian kontekstual yaitu harus mempertimbangkan kebutuhan dan kondisi guru termasuk beban kerja guru pada unit kerja masing-masing. Kebutuhan penguasaan asesmen guru di sekolah dasar mungkin berbeda dengan kebutuhan guru di level sekolah menengah.

Tidak dipungkiri memang keterbatasan ilmu atau keterampilan tentang literasi penilaian bahasa sangat banyak dialami oleh para guru. Minimnya sumber belajar dan kesempatan untuk mengikuti pelatihan menjadi faktor utama yang selalu menjadi alasan. Maka dari itu, buku ini akan memberi pencerahan tentang prinsip-prinsip penilaian dan pengajaran, teknik dalam penilaian, apa saja yang harus dinilai, serta bagaimana cara pengambilan nilai dalam penilaian bahasa yang akan dibahas pada bab-bab berikutnya.

Bab 2

Prinsip-prinsip Penilaian

Objek Pembelajaran

- A. Praktis
- B. Andal/Reliabel
- C. Sahih/Valid
- D. Autentik
- E. Washback

Pada bab ini, akan dijelaskan secara detail mengenai prinsip-prinsip penilaian khususnya dalam penilaian bahasa (*language assessment*). Bab ini akan mengeksplor tentang keefektifan sebuah tes dengan mempertimbangkan prinsip-prinsip yang ada dan bagaimana menyeimbangkan prinsip-prinsip ini dalam praktik penilaian di kelas sehari-hari.

A. Praktis

Prinsip pertama adalah kepraktisan. Tidak diragukan lagi bahwa tes yang baik atau bisa dikatakan tes itu efektif apabila tes tersebut praktis. Artinya, dalam mengadakan sebuah tes, tes tersebut tidak harus mengeluarkan banyak biaya, sesuai batas waktu yang disediakan, mudah untuk digunakan, dan memiliki prosedur penilaian yang spesifik (Brown, 2004).

Tes yang mahal biayanya tidak praktis. Ujian kemampuan berbahasa yang membutuhkan siswa lima jam untuk menyelesaikannya tidak praktis--tes ini menghabiskan lebih banyak waktu (dan uang) daripada yang diperlukan untuk mencapai tujuannya. Sebuah tes yang dilakukan oleh individu secara satu per satu tidak praktis untuk kelompok yang terdiri dari beberapa ratus siswa dan hanya sedikit guru yang melakukan pengujian. Contoh lain adalah jika sebuah tes yang hanya memakan waktu beberapa menit bagi seorang siswa untuk mengerjakannya namun perlu beberapa jam bagi guru untuk menilainya bukanlah tes yang praktis.

Bachman & Palmer (1996) menetapkan rumus khusus untuk mengukur kepraktisan sebuah tes. Untuk menilai kepraktisan, kita perlu menentukan apa yang kita maksud dengan sumber daya. Bachman & Palmer mengklasifikasikan sumber daya ini menjadi tiga tipe, yaitu (1) sumber daya manusia, termasuk penulis ujian, pencetak gol atau penilai, dan administrator pengujian, serta personel pendukung administrasi dan teknis; (2) sumber daya material termasuk ruang, peralatan (misalnya komputer atau video) dan bahan (misalnya kertas dan gambar), serta (3) waktu yang meliputi waktu untuk pengembangan soal, lamanya tes berlangsung, waktu penilaian atau penskoran. Dalam mempertimbangkan berbagai jenis sumber daya, penting juga untuk memperkirakan biaya (keuangan) yang mungkin terkait dengan masing-masing sumber daya, karena biaya berkaitan erat dengan kemampuan pengadaan sumber daya.

Jadi, kepraktisan sebuah tes dapat diukur dari sejauh mana tuntutan spesifikasi tes tertentu dapat dipenuhi dengan sumber daya yang ada. Jika sumber daya yang dibutuhkan tidak melebihi sumber daya yang tersedia pada setiap tahap dalam pengembangan tes, maka tes dapat dikatakan praktis dan pengembangan dan penggunaan tes dapat dilanjutkan. Jika sumber daya yang tersedia kurang, maka tes tidak

praktis dan pengembang tes atau guru harus memodifikasi spesifikasi untuk mengurangi sumber daya yang diperlukan, atau menambah sumber daya yang tersedia sehingga dapat digunakan lebih efisien. Dengan demikian, tes praktis adalah tes yang desain, pengembangan, dan penggunaannya tidak memerlukan lebih banyak sumber daya yang tersedia.

B. Andal/Reliabel

Reliabilitas sebuah tes merupakan hal yang harus dipertimbangkan oleh pengembang/pembuat tes untuk melihat apakah tes yang dibuat konsisten dan dapat dipertanggungjawabkan. Tes yang reliabel itu merupakan tes yang jika diujikan kepada siswa yang sama di waktu yang berbeda maka jawaban atau hasil dari siswa hendaknya akan sama juga dari tes sebelumnya. Isi dari tes atau soal yang tidak ambigu pasti akan menghasilkan jawaban yang pasti, maksudnya jika soal atau tes tersebut diulang di waktu yang berbeda maka jawabannya tetap sama.

Keandalan sering didefinisikan sebagai konsistensi pengukuran. skor tes yang andal akan konsisten di berbagai karakteristik situasi pengujian. Oleh karena itu, keandalan dapat dianggap sebagai fungsi dari konsistensi skor dari satu set tes dan tugas-tugas tes yang lain adalah kita menganggap tes sebagai sekumpulan karakteristik tugas, maka reliabilitas dapat dianggap sebagai fungsi konsistensi antar sekumpulan karakteristik tugas yang berbeda.

Dengan demikian, dalam tes yang dirancang untuk mengubah urutan individu dari tertinggi ke terendah, jika skor yang diperoleh pada bentuk yang berbeda tidak memberi peringkat individu pada urutan yang sama, maka skor ini sangat tidak konsisten dan akan dianggap sebagai indikator kemampuan yang tidak dapat dilemahkan.

Keandalan jelas merupakan kualitas penting dari nilai tes, karena kecuali jika nilai tes konsisten, mereka tidak dapat memberikan kita informasi apa pun tentang kemampuan yang ingin kita ukur. Pada saat yang sama kita perlu menyadari bahwa tidak mungkin untuk menghilangkan ketidakkonsistenan sepenuhnya. Apa yang dapat kita lakukan, bagaimanapun adalah mencoba untuk meminimalkan efek dari sumber-sumber potensial yang tidak konsisten yang berada di bawah kendali kita, melalui desain uji.

Brown (2004) membagi lagi prinsip keandalan ini menjadi lebih spesifik seperti di bawah ini.

1. *Student-Related Reliability*

Dalam faktor yang satu ini, ketidakreliabilisan sebuah tes terjadi akibat masalah-masalah yang dialami siswa. Seperti contoh siswa mengalami sakit, kelelahan, kecemasan, dan banyak masalah fisik atau psikologi yang dialami siswa pada saat mengerjakan tes. Tentu saja jika tes yang sama namun diulang dalam waktu yang berbeda faktor ini sangat mungkin terjadi dan bisa dialami oleh setiap siswa. Kondisi ini akan memengaruhi hasil tes dan tidak mencerminkan kemampuan sesungguhnya.

2. *Rater-Related Reliability*

Dalam pemberian skor ada dua kemungkinan masalah ini bisa terjadi. *Pertama, inter-rater reliability*. Dalam masalah ini ketidakreliabilisan bisa terjadi ketika tes yang sama diujikan dalam waktu yang berbeda namun ada dua penguji yang berbeda yang mengadakan tes tersebut. Mengapa demikian? Sudah pasti, karena penguji satu dengan yang lainnya bisa jadi memiliki kriteria yang berbeda dalam menentukan skor siswa. Perbedaan pengalaman, perbedaan pemahaman terhadap proses penilaian, dan bisa menjadi tolok ukur dalam menentukan reliabel tidaknya sebuah tes. Selanjutnya, ada masalah *intrarater reliability* yang berkaitan dengan individu pemberi skor atau guru itu sendiri. Misalnya karena faktor kelelahan saat menilai ujian siswa dalam jumlah banyak, guru bisa saja memberi nilai yang tidak adil kepada siswa yang diperiksa lebih awal atau lebih akhir. Masalah berkaitan dengan suka dan tidak sukarnya guru terhadap siswa tertentu bisa juga memengaruhi keandalan tes dilihat dari sisi *intra-rater reliability*.

3. *Reliabilitas Administrasi Tes*

Dalam pelaksanaan sebuah tes, terkadang situasinya tidak sesuai dengan apa yang kita harapkan. Katakanlah ketika kita mengadakan sebuah tes online. Pada tes sebelumnya mungkin hasil dari tes lumayan menunjukkan angka yang bagus. Namun, pada tes yang sama dilaksanakan pada kesempatan yang berbeda, hasilnya sangat berbeda

dari yang sebelumnya. Ternyata pada tes yang kedua ada kendala koneksi internet yang memengaruhi hasil dari tes tersebut. Ketidakstabilan jaringan internet tersebut menjadi tolok ukur bahwa pelaksanaan tes juga dapat menimbulkan ketidakreliabilisan dari tes tersebut. Contoh lain adalah kondisi ruang kelas yang berbeda saat melakukan tes yang pertama dan yang kedua. Misalnya ruang kelas yang bising dan panas akan memengaruhi performa siswa.

4. Reliabilitas Tes

Terkadang tes yang sudah kita buat juga mempunyai kekurangan-kekurangan yang mungkin dapat menimbulkan masalah atas keandalannya. Soal yang terlalu panjang, atau soal yang ambigu merupakan hal yang harus dihindari dalam masalah ini. Selain kedua aspek tersebut, batas waktu pada saat tes juga penting karena kebanyakan siswa mengeluh ketika waktu yang diberikan habis. Tentu dalam hal ini waktu ujian yang tidak sesuai dengan banyaknya soal yang diujikan. Brown (2004) menegaskan bahwa soal esai yang ambigu merupakan sumber utama tidak reliabilitasnya sebuah tes.

C. Sahih/Valid

Validitas atau kesahihan merupakan salah satu prinsip yang paling penting dalam pelaksanaan sebuah tes. Mengapa dikatakan penting, karena tes yang valid merupakan tes yang sesuai dengan tujuan pembelajaran atau materi yang diajarkan. Dikatakan dalam bukunya, Brown (2004) dikutip dari Gronlund (1998) mengatakan bahwa tes yang valid bisa dilihat jika hasil dari sebuah penilaian itu sesuai, bermakna, dan bermanfaat dengan tujuan dilaksanakannya penilaian tersebut. Selain itu, Bachman & Palmer (1996) mengungkapkan bahwa investigasi validitas berkenaan dengan pertanyaan, 'Sejauh mana hasil tes seseorang terkait dengan kemampuan berbahasa yang akan diukur?'

Implikasi yang jelas dari pertanyaan ini adalah bahwa sebagai pengembang tes atau guru, kita harus dapat memberikan justifikasi yang memadai dari skor tes yang diberikan. Artinya, kita perlu menunjukkan, atau membenarkan, validitas interpretasi yang kita buat terhadap nilai ujian, dan tidak sekadar menyatakan atau berpendapat bahwa itu valid. Sebelum guru melaksanakan sebuah tes, alangkah baiknya harus

mempertimbangkan terlebih dahulu jenis-jenis validitas. Menurut Brown (2004), ada lima jenis validitas yang harus diperhatikan.

1. *Content Validity*

Validitas ini sangat berhubungan erat dengan isi dari sebuah tes. Sebagai contoh ketika ada ulangan menulis (*writing*) mengenai naratif, maka isi soal yang akan muncul juga seputar cerita-cerita dongeng, legenda, dan cerita fiktif lainnya. Contoh kasus tersebut menggambarkan bahwa soal ulangan tersebut memiliki validitas isi yang tinggi, karena ada kesesuaian antara isi soal dengan materi atau tujuan pembelajaran yang diujikan. Contoh lain juga diungkapkan oleh Brown (2004) jika dalam sebuah materi ada lima tujuan pembelajaran yang harus dicapai, namun hanya sebagian saja yang diujikan dalam tes, maka validitas isi dari tes tersebut bisa dibilang rendah. Dari kedua contoh di atas dapat disimpulkan bahwa *content validity* atau validitas isi merupakan tingkat relevansi antara isi soal dengan materi dan tujuan pembelajaran.

2. Validitas Kriteria

Jenis validitas yang kedua yaitu validitas kriteria atau *criterion-related validity*. Pada validitas jenis ini menurut Brown (2004) guru mendesain sebuah tes harus sesuai dengan konsep dari kriteria-kriteria penilaian. Atau dengan kata lain bisa diartikan bahwa skor yang dihasilkan pada saat tes mampu mengukur kriteria dari tujuan pembelajaran. Brown juga menambahkan bahwa validitas kriteria dapat dicapai melalui perbandingan hasil penilaian satu dengan yang lainnya asalkan masih dalam satu kriteria. Hal tersebut dapat digambarkan ketika guru hendak menilai kemampuan tata bahasa siswa menggunakan metode *error analysis* pada *reading text*. Untuk mendapatkan validitas kriteria dari tes tersebut guru ternyata membandingkan secara bebas (bukan tes yang dibuat oleh guru tersebut) dengan hasil skor dari *writing test* di mana murid harus menulis dengan *grammar* yang tepat. Dari contoh di atas cara yang dilakukan guru untuk mencari tingkat validitas kriteria cukup efektif. Serupa dengan contoh di atas Brown (2004) juga menggambarkan situasi di mana guru ingin melakukan tes secara verbal mengenai *phonetic voice* dan *voiceless*. Hasil dari tes tersebut kemudian dibandingkan secara bebas (dengan guru lain) dengan hasil tes *phonetic* yang dilakukan menggunakan *textbook*. Dapat kita lihat dari

kedua contoh di atas bahwa validitas kriteria dapat diukur dengan cara membandingkan hasil dari skor tes yang di-*treatment* berbeda namun masih dalam satu kriteria tujuan atau materi pembelajaran.

3. Validitas Konstruk

Untuk menjustifikasi interpretasi skor tertentu, kita perlu memberikan bukti bahwa skor tes mencerminkan area kemampuan yang ingin kita ukur. Untuk memberikan bukti seperti itu, kita harus mendefinisikan konstruk yang ingin kita ukur. Bachman & Palmer (1996) menyatakan bahwa validitas konstruk adalah definisi dari kemampuan yang memberikan dasar untuk menafsirkan skor tes. Oleh karena itu, istilah validitas konstruk digunakan untuk merujuk pada sejauh mana kita dapat menginterpretasikan skor tes yang diberikan sebagai indikator kemampuan, atau konstruk yang ingin kita ukur. Validasi konstruk adalah proses yang berlangsung terus-menerus untuk menunjukkan bahwa interpretasi dari nilai tes didukung oleh bukti-bukti. Penting bagi pengembang tes, guru maupun pengguna tes untuk menyadari bahwa validasi tes adalah proses yang berkelanjutan dan bahwa interpretasi yang kita lakukan terhadap skor tes tidak pernah dapat dianggap benar-benar valid. Justifikasi terhadap interpretasi yang kita buat berdasarkan skor tes dimulai dengan desain tes dan berlanjut dengan pengumpulan bukti untuk mendukung interpretasi yang kita maksudkan. Bagi guru kelas, prosedur validasi konstruk mungkin terlihat sebagai proses yang merepotkan. Namun demikian, melakukan validasi konstruk untuk setiap tes yang digunakan di kelas sangat penting dan layak. Misalnya, sebagai guru, Anda melakukan tes kosakata dalam bentuk pilihan ganda. Format tes seperti ini mungkin sudah mencapai pembahasan yang dicakup dalam unit pembelajaran. Namun demikian, jika tujuan dari pembelajaran tersebut adalah menggunakan kosakata secara komunikatif, penggunaan format pilihan ganda tidak sesuai dengan definisi konstruk penggunaan bahasa komunikatif.

4. Validitas Konsekuensial

Berbeda dengan ketiga jenis validitas sebelumnya, validitas konsekuensial ini merupakan validitas pilihan, karena hanya guru-guru yang mempunyai tujuan dan keinginan tertentu dalam memvalidasi tesnya yang memakai jenis validitas ini. Menurut Brown (2004), validitas konsekuensial

ini mencakupi seluruh konsekuensi ujian, termasuk pertimbangan keakuratan dalam mengukur kriteria yang dimaksud, dampak pada persiapan tes yang diuji, dampak terhadap siswa yang melakukan tes, dan konsekuensi sosial dari interpretasi dan penggunaan tes.

5. *Face Validity*

Face validity merupakan salah satu aspek penting dari validitas konsekuensial yang mana siswa memandang bahwa penilaiannya adil, relevan, dan berguna untuk meningkatkan pembelajaran (Groundlound, 1998). Menurut Brown (2004), tingkat *face validity* akan tinggi atau terpenuhi jika siswa menjumpai seperti hal di bawah ini.

- a. Konstruksi tes nya bagus, sesuai dengan format tes yang familiar.
- b. Sebuah tes yang jelas dapat dilakukan dalam batas waktu yang diberikan.
- c. Soalnya jelas dan tidak berbelit-belit.
- d. Petunjuk tesnya jelas.
- e. Soalnya berhubungan dengan materi belajar.
- f. Tingkat kesulitannya masih dalam tingkat yang wajar.

Hal yang perlu diingat bahwa *face validity* bukan sesuatu yang dapat diukur secara empiris oleh guru atau bahkan ahli testing. Hal ini semata-mata merupakan faktor secara intuisi yang dapat dirasakan siswa atau si pemberi tes.

Tujuan utama tes adalah sebagai alat ukur yang dapat kita jadikan sebagai indikator kemampuan seseorang. Dua prinsip penilaian, yaitu reliabilitas dan validitas dengan demikian penting untuk digunakan dalam tes apa pun. Reliabilitas adalah kondisi yang diperlukan untuk validitas. Namun reliabilitas bukanlah kondisi yang memadai untuk membangun validitas. Sebagai contoh, sebuah perusahaan memerlukan tes untuk menempatkan individu ke level yang berbeda dalam kemampuan menulis. Tes pilihan ganda tentang kosakata mungkin menghasilkan skor yang sangat konsisten atau andal, tetapi ini tidak akan cukup untuk membenarkan menggunakan tes ini sebagai tes penempatan untuk kelas menulis. Ini dikarenakan pengetahuan kosakata hanya satu aspek dari kemampuan menggunakan bahasa untuk melakukan tugas-tugas penulisan akademik.

D. Autentik

Prinsip keempat dalam penilaian bahasa adalah *authenticity*. Autentik atau keaslian berkaitan dengan kesesuaian tugas penilaian bahasa dengan situasi komunikasi yang sesungguhnya atau dunia nyata. Penilaian autentik, di antaranya penulisan esai, laporan, desain proposal, portofolio (Mcnamara, 2009). Menurut Brown (2004) penilaian atau tes dapat dikatakan autentik jika (1) bahasa yang digunakan pada soal-soal merupakan bahasa yang alami atau wajar untuk digunakan; (2) soal-soalnya harus kontekstual dan tidak terisolasi; (3) topik yang digunakan harus bermakna (relevan, menarik) bagi siswa; (4) rangkaian-rangkaian soal membentuk cerita atau episode; (5) tugas-tugas penilaian mendekati tugas-tugas dalam kehidupan nyata.

Untuk mewujudkan tes yang autentik memang gampang-gampang susah. Guru diharapkan untuk bekerja dengan kreatif; menghubungkan soal yang akan diujikan dengan realita yang ada. Misalnya, guru hendak mengadakan tes *reading comprehension*. Pada saat membuat soal guru harus mencari referensi yang sedang tren di kalangan siswa, seperti *bullying*, prestasi murid yang mengikuti kejuaraan internasional dan sebagainya. Memang sangat bagus jika setiap guru menerapkan seperti demikian. Tentu hal ini akan membuat siswa lebih semangat dan tentunya dia tidak cepat bosan terutama pada saat mengerjakan tes.

E. Washback

Dalam bukunya, Brown (2004: 28) menyatakan bahwa dalam penilaian bahasa *washback* merupakan pengaruh tes dalam proses belajar mengajar. Efek *washback* sendiri terbagi menjadi dua macam, yaitu *positive washback* dan *negative washback*. Salah satu cara memberikan *positive washback*, yaitu membangun ilmu pengetahuan pada siswa dengan cara memberikan aspek-aspek penunjang demi keutuhan dari pengajaran. Maksudnya, dengan cara memberi *feedback* kepada siswa tentang kekuatan dan kelemahan mereka tentunya setelah mereka melaksanakan tes. Cara ini dapat meningkatkan efek *washback* pada pembelajaran. Contohnya dengan memberikan pujian kepada yang nilainya baik dan komentar yang membangun bila hasilnya kurang baik.

Berbeda dengan *positive washback*, *negative washback* juga mempunyai cara sendiri dalam memberikan pengaruh dalam proses belajar mengajar.

Negative washback salah satunya ditandai dengan perencanaan kurikulum yang menitikberatkan hanya pada satu aspek keterampilan. Jika dalam *positive washback* itu *treatment*-nya muncul setelah diadakannya sebuah tes, berbeda dengan *negative washback* yang tentunya efek ini terjadi sebelum dimulainya sebuah tes. Seperti yang dinyatakan oleh Brown (2004: 29), Beliau menegaskan bahwa *washback* juga termasuk efek dari sebuah penilaian dalam belajar mengajar sebelum penilaian tersebut dilaksanakan. Hal tersebut bisa digambarkan pada saat akan dilaksanakannya ujian akhir bahkan UN. Hampir seluruh instansi atau guru-guru mempraktikkan efek tersebut. Adanya bimbel (Bimbingan Belajar) secara intens dilakukan agar para siswa memenuhi ketentuan kelulusan nilai pada ujian akhir. Sebaliknya, salah satu cara untuk meningkatkan efek *washback* adalah dengan cara memberi komentar kepada setiap hasil ulangan siswa secara spesifik (Brown, 2004: 29).

Motivasi ekstrinsik mungkin lebih baik daripada tidak ada motivasi sama sekali--maka setiap tes, baik atau buruk, dapat dikatakan mengalami kemunduran yang bermanfaat jika meningkatkan aktivitas atau motivasi tersebut.

Latihan

Untuk mengetahui sejauh mana pemahaman Anda tentang materi yang dipaparkan dalam bab ini, kerjakanlah latihan berikut ini.

1. Tinjau lima prinsip dasar penilaian yang didefinisikan dan dijelaskan dalam bab ini. Pastikan untuk membedakan antara beberapa jenis bukti yang mendukung validitas suatu tes, serta empat jenis reliabilitas!
2. Menurut Anda apakah konsekuensi dan validitas wajah merupakan pertimbangan yang tepat dalam penilaian berbasis kelas? Jelaskan!
3. *Washback* dijelaskan di sini sebagai dampak positif tes. Dapatkah tes memberikan *washback* negatif? Jelaskan!

Bab 3

Penilaian dan Pengajaran

Objek Pembelajaran

- A. Penilaian Formal dan Informal
- B. Penilaian Formatif dan Sumatif
- C. *Norm-Referenced* dan *Criterion-Referenced Testing*

Pada bab ini akan dibahas istilah dari penilaian dan pengajaran. Di dalam proses belajar dan pembelajaran pastilah tidak pernah terpisahkan dengan istilah penilaian dan pengajaran atau *assessment and teaching*. Mungkin tanpa adanya kedua istilah tersebut proses pembelajaran tidak akan berjalan dengan baik. Di sisi lain, banyak orang beranggapan bahwa istilah *assessment* itu sama dengan *test*. Seperti yang dinyatakan oleh Brown (2004: 4) bahwa kebanyakan orang berpikir bahwa kata asesmen dan tes itu adalah sebuah sinonim. Namun pada kenyataannya kedua istilah tersebut sangat berlainan makna. Maka dari itu, pada kesempatan kali ini kita akan jelaskan lebih lanjut apa itu tes, asesmen, dan pengajaran.

Apakah penilaian merupakan bagian dari pengajaran? Atau apakah mungkin tes adalah bagian terkecil dari sebuah pengajaran? Mari kita bahas satu per satu untuk ketiga istilah tersebut. *Teaching* atau pengajaran merupakan kegiatan dalam proses belajar mengajar di mana guru memberikan ilmunya dan memberikan penilaian kepada siswa. Saat mengajar, guru memberi kesempatan kepada siswa untuk berpikir, mendengarkan, menentukan tujuan, memproses umpan balik dari guru, yang mana serangkaian aktivitas tersebut akan mengasah kemampuan siswa dalam memahami apa yang harus dipahami. Untuk melihat sejauh mana pemahaman siswa mengenai materi yang telah sedang diajarkan guru, guru melakukan aktivitas yang bernama asesmen atau penilaian. Sedangkan tes itu sendiri merupakan bagian dari asesmen (Brown, 2004) yang digunakan untuk mengukur pemahaman siswa terhadap materi yang diajarkan oleh guru.

A. Penilaian Formal dan Informal

Asesmen formal segala sesuatunya sudah diset dan dipersiapkan dengan baik oleh guru. Asesmen formal tidak bersifat spontan atau *incidental*. Jenis asesmen ini biasanya digunakan secara periodik dalam sebuah proses pembelajaran. Apakah asesmen formal sama dengan tes? Semua tes adalah bentuk asesmen formal, namun tidak semua asesmen formal adalah tes. Contohnya, ulangan harian atau ulangan akhir semester. Selain itu, dalam asesmen formal juga tidak melulu tentang tes. Brown (2004) menyatakan bahwa jurnal dan portofolio siswa juga dapat dijadikan umpan balik bagi siswa untuk melihat pencapaian dari sebuah objek pembelajaran.

Berbeda dengan asesmen formal, jenis asesmen informal ini sering dilakukan di kelas. Kita ambil contoh ketika siswa berhasil menjawab pertanyaan guru, lalu guru membalasnya dengan “bagus sekali”. Ungkapan tersebut sudah merupakan penerapan sebuah asesmen, dan hal ini biasanya terjadi secara spontan atau *incidental*. Asesmen informal memiliki banyak bentuk, seperti melakukan komentar yang memang tidak direncanakan, pujian, dan saran yang memang tidak disengaja. Pastinya, jika kita lihat karakteristiknya, jenis penilaian ini dilakukan ketika pelajaran berlangsung. Dalam asesmen informal, umpan baliknya tidak diberikan melalui skor hasil tes, tapi dari hasil diskusi, dan bisa jadi dari komentar-komentar penampilan siswa. Contoh lain ketika

guru memberi emoji dalam memberikan umpan baliknya pada lembar pekerjaan siswa. Selain memang penilaian ini diberi secara personal kepada siswa atau tidak melibatkan siswa lain sebagai pembanding, kadang ucapan-ucapan seperti *nice job* dan lain-lain, mengalir begitu saja tanpa sadar bahwa sebenarnya guru tengah memberikan sebuah asesmen yang sifatnya informal. Ternyata hal tersebut ditegaskan oleh Brown (2004) yang menyatakan bahwa asesmen informal juga terdapat di dalam tugas-tugas siswa di mana tidak ada unsur untuk menjustifikasi kemampuan siswa. Contohnya, ketika guru memberikan komentar di pinggir dari esai yang ditulis siswa dan memberi masukan bagaimana cara menulis kalimat yang baik dalam Bahasa Inggris.

B. Penilaian Formatif dan Sumatif

Cara lain untuk mengingat perbedaan asesmen dan tes, yaitu dengan melihat fungsi dari asesmen. Penilaian sumatif dan formatif merupakan dua fungsi yang berbeda untuk sebuah asesmen. Fungsi dari penilaian formatif bekerja pada jenis penilaian informal. Berarti mengevaluasi siswa dalam proses “*forming*” atau pembentukan kompetensi dan *skill* mereka dengan tujuan untuk membantu siswa dalam meraih tujuan pembelajaran.

Dengan demikian, ketika misalnya guru mengajar Bahasa Inggris di kelas, guru bisa memberikan komentar kepada siswa, memberi saran bagaimana mengucapkan *pronunciation* yang benar, atau mengingatkan jika siswa melakukan kesalahan dalam menggunakan tata bahasa. Umpan balik seperti itulah yang digunakan di dalam penilaian formatif. Tentunya, sebelum umpan balik tersebut diberikan biasanya ada latihan atau tugas yang harus dikerjakan siswa seperti presentasi, diskusi, kuis, dan lain-lain. PR (pekerjaan rumah) juga merupakan bentuk dari penilaian formatif. Mengapa? Karena PR diberikan berdasarkan materi yang sedang diajarkan, jadi secara tidak langsung guru sedang melakukan observasi untuk mencari tahu sudah sejauh mana kemampuan siswanya.

Berbeda dengan penilaian formatif, fungsi dari penilaian sumatif juga untuk mengukur. Maksudnya, setelah siswa dirasa cukup menguasai materi yang telah diberikan oleh guru, saatnya guru mengukur kemampuan siswanya. Contoh penilaian sumatif di antaranya seperti ujian akhir, ulangan semester, ujian kenaikan kelas, dan lain-lain.

Memang untuk sebagian banyak dari siswa, tes merupakan sesuatu yang menakutkan bahkan bisa menimbulkan *anxiety* (kecemasan).

Hal tersebut merupakan masalah yang sudah umum terjadi di dunia pendidikan. Tugas guru di sini untuk merubah paradigma tersebut mungkin dengan cara yang lebih menyenangkan ketika mengadakan sebuah tes untuk meningkatkan motivasi siswa tersebut.

C. *Norm-Referenced dan Criterion-Referenced Testing*

Satu lagi hal yang penting untuk mengklarifikasi perdebatan antara asesmen dan tes yaitu mengenai perbedaan antara *norm-referenced* (Penilaian Acuan Normatif) dan *criterion-referenced testing* (Penilaian Acuan Patokan). Di dalam *norm-referenced testing*, tes ini ditujukan kepada siswa untuk menempatkannya ke dalam sebuah *ranking* (urutan) di dalam suatu kelompok (Brown, 2004: 7). Maksudnya, jenis penilaian ini tidak digunakan untuk mengukur kemampuan siswa dalam penguasaan materi pembelajarannya, melainkan hanya membandingkan penguasaan individu terhadap rata-rata penguasaan kelompok. Masih bingung? Anda tahu tes IQ? Iya, tes IQ merupakan salah satu dari sekian banyak tes yang masuk dalam golongan *norm-referenced testing*. Seperti yang kita ketahui bahwa tujuan tes IQ tidak untuk mengukur kemampuan siswa dalam belajar, tapi dia lebih membandingkan hasil tes siswa satu dengan yang lain. Contoh lain yaitu *placement test* atau tes penempatan. Sama seperti tes IQ prinsip dan tujuannya sama, yaitu untuk membandingkan nilai para siswa sesuai dengan kriterianya. Hughes (1993) juga memberi contoh jika jenis penilaian ini digunakan untuk mengklasifikasikan nilai siswa berdasarkan tinggi-rendahnya, Beliau menyatakan bahwa nilai siswa akan menempatkan mereka di sepuluh teratas dari keseluruhan siswa atau siswa juga bisa lima terbawah dari keseluruhan siswa yang mengikuti tes tersebut. Contoh lain jika ada siswa yang mendapat skor di atas 80 pada *placement test*-nya akan ditempatkan pada kelas unggulan dalam sekolahnya, jika siswa mendapat nilai yang jelek maka akan menempati kelas biasa. Seperti itu kira-kira ilustrasi mengenai *norm-referenced testing*.

Di sisi lain, ada *criterion-referenced testing* atau Penilaian Acuan Patokan. Jenis penilaian ini berbanding terbalik dengan *norm-referenced testing*. Mengapa demikian? Seperti yang kita tahu dalam *norm-referenced testing*, siswa tidak mendapatkan *feedback* sama-sekali. Pada penilaian ini justru ditujukan untuk memberikan *feedback* kepada siswa, yang biasanya *feedback* tersebut berupa angka atau skor dari tes yang mengukur

sejauh mana kemampuan belajar siswa. Pandangan yang hampir sama mengenai tujuan dari diadakannya penilaian ini yaitu untuk mengklasifikasikan nilai siswa berdasarkan kemampuannya dalam mengerjakan tugas atau tes yang sesuai dengan tujuan pembelajaran. Dalam tes jenis ini, siswa dituntut untuk mempertanggungjawabkan apa yang telah mereka dapatkan atau pahami selama proses pembelajaran. Apakah semuanya sudah sesuai dengan kriteria (silabus, kompetensi dasar, dan lain-lain?) Nah tentu hasil dari tes ini mempunyai umpan balik, tidak seperti *norm-referenced testing*. Umpan balik yang dilakukan bertujuan untuk mengetahui apakah siswa mereka benar-benar sudah menguasai pembelajaran yang sudah diajarkan atau belum.

Latihan

Kerjakan soal-soal berikut ini untuk mengukur pemahaman Anda tentang isi bab ini.

1. Mengapa portofolio disebut dengan penilaian formatif?
2. Apa saja yang manfaat dari melaksanakan penilaian formatif?
3. Bisakah fungsi penilaian informal menggantikan penilaian formal?

[Halaman ini sengaja dikosongkan]

Bab 4

Penilaian Kelas

Objek Pembelajaran

- A. Penilaian Pembelajaran (*Assessment of Learning*)
- B. Penilaian untuk Pembelajaran (*Assessment for Learning*)
- C. Penilaian Sebagai Pembelajaran (*Assessment as Learning*)

Sebelum kita membahas lebih jauh mengenai bab ini, mari kita lihat apakah ada perbedaan antara standardisasi penilaian di kelas dan standardisasi penilaian dengan skala yang lebih besar (Fulcher, 2010: 67). Tidak dipungkiri kedua jenis asesmen di atas menjadi hal yang diperdebatkan oleh para ahli. Namun pada kenyataannya kedua asesmen tersebut memang harus saling berkaitan satu sama lain.

Bisa dilihat ketika guru sudah mendesain sendiri untuk penilaian sesuai keadaan di kelasnya, tidak dapat dipungkiri akan selalu ada tempat di mana standardisasi atau kriteria dalam penilaian (*standardized assessment*) dijadikan acuan untuk melaksanakan penilaian di kelas. Hal ini terjadi karena tes bukan sekadar alat yang kita pakai untuk mengukur apa yang hendak kita ukur dalam kegiatan belajar mengajar (Fulcher, 2010). Beliau menambahkan bahwa tes yang sesuai standar itu juga berfungsi untuk meningkatkan motivasi siswa dalam belajar. Maka dari itu, idealnya, penilaian kelas memang harus disesuaikan dengan standar atau kriteria dari sebuah penilaian.

Tidak dapat dipungkiri jika sebuah nilai dapat dijadikan mesin untuk mendorong pembelajaran siswa. Namun alangkah lebih baiknya jika tidak hanya nilai yang dijadikan tolok ukur untuk mendorong kegiatan belajar mengajar di kelas. Dalam penilaian di kelas, kita harus menciptakan mesin pendorong untuk meningkatkan sebuah pembelajaran, tidak hanya dengan cara memotivasi siswa, tetapi juga dengan memberikan umpan pada pembelajaran dan pencapaian dari siswa maupun guru.

Pada bab ini kita akan membahas tiga pendekatan utama dalam penilaian di kelas. Pertama kita akan membahas pendekatan pertama, yaitu mengenai *assessment of learning* (penilaian pembelajaran), selanjutnya kita akan membahas tentang *assessment for learning* (penilaian untuk pembelajaran), dan yang terakhir yaitu *assessment as learning* (penilaian sebagai pembelajaran).

A. Penilaian Pembelajaran (*Assessment of Learning*)

Pendekatan ini merupakan pendekatan yang paling populer digunakan di kalangan para pengajar atau guru dalam memberikan laporan hasil pembelajaran siswa. Bagaimana tidak, jika pendekatan ini tidak digunakan ketika guru mengajar di kelas, maka siswa bahkan orang tua siswa tidak akan tahu kemajuan kemampuan yang sudah dicapai. Seperti yang diungkapkan Lorna (2003), bahwa tujuan dari penilaian pembelajaran itu untuk memberi tahu kepada siswa maupun orang tua mengenai kemajuan pembelajaran yang ada di sekolah, biasanya dilakukan dengan membandingkan posisi dari satu siswa dengan lainnya atau dikenal sebagai sistem *ranking*. Biasanya penilaian pembelajaran akan diberikan pada akhir dari pembelajaran, materi belajar, bahkan akhir semester dalam proses pembelajaran. Untuk melihat kemajuan atau hasil belajar siswa, dalam pendekatan ini menggunakan tes atau ujian yang berisi pertanyaan seputar materi yang telah diajarkan selama proses belajar.

Jenis penilaian seperti ini sangat banyak digunakan untuk penilaian kelas terutama untuk sekolah menengah. Di Indonesia pun untuk sistem penilaian seperti ini banyak sekali digunakan terutama di SMP, SMA, bahkan perguruan tinggi sekalipun. Dalam penilaian pembelajaran, guru kerap kali menilai pencapaian kompetensi dasar secara kuantitas

dengan menggunakan simbol atau angka-angka untuk menunjukkan bobot dari hasil pencapaian siswa. Angka atau simbol tersebut dijadikan sebuah umpan balik untuk siswa di mana sangat minim sekali informasi berupa arahan atau masukan mengenai pembelajaran siswa yang sudah dipelajari (Lorna, 2003). Tentunya peran guru pada pendekatan jenis ini begitu krusial. Selain mempersiapkan tes untuk melihat prestasi siswa, guru juga dituntut harus adil dan transparan pada saat memberikan skor.

Dengan melihat fakta yang ada, nyatanya pendekatan penilaian jenis ini dapat diterima oleh masyarakat secara umum; siswa dan orang tua. Tidak dipungkiri banyak sekali pro dan kontra di kalangan para pakar pendidikan mengenai keadilan, transparansi, dan akurasi dalam penentuan skor kepada siswa. Para peneliti pendidikan dan ahli teori telah mengkaji mengenai sistem penskoran tradisional ini dalam kurun beberapa waktu (Lorna, 2003; Marzano, 2000). Sementara itu, masalah yang timbul dalam kajian tersebut yaitu mengenai perbedaan bobot penilaian antara satu guru dengan yang lainnya, salah penafsiran skor pada penilaian yang mewakili bidang keterampilan dan kemampuan siswa (Marzano, 2000).

B. Penilaian untuk Pembelajaran (*Assessment for Learning*)

Jika pada pendekatan sebelumnya kita membahas tentang penilaian pembelajaran yang fokus pada penilaian sumatif di mana proses evaluasi dilakukan pada akhir dari sebuah pembelajaran, pada pendekatan kali ini kita akan menjelaskan apa itu penilaian untuk pembelajaran *assessment for learning*. Pada pendekatan penilaian yang satu ini jenis penilaiannya fokus terhadap penilaian formatif di mana proses evaluasi berjalan beriringan dengan proses pembelajaran. Bahkan Fulcher (2010: 68) menyatakan bahwa akhir-akhir ini penilaian digunakan saat proses pembelajaran berlangsung untuk meningkatkan pembelajaran itu sendiri, daripada melaksanakannya di akhir dari pembelajaran. Sama dengan apa yang Fulcher sampaikan, Lorna (2003) juga menyatakan bahwa penilaian untuk pembelajaran itu terjadi di tengah atau pada saat pembelajaran berlangsung, bukan di akhir dari sebuah pembelajaran (*summative*), dan biasanya lebih dari sekali pelaksanaannya. Tidak seperti tes yang biasa diadakan, pada pendekatan ini sebuah evaluasi ditujukan untuk mendiagnosa kekurangan dan kelebihan siswa pada saat proses belajar berlangsung.

Dalam penilaian jenis ini, umpan balik yang di dalamnya berisi masukan atau saran lebih diutamakan daripada memberi skor kepada siswa (Fulcher, 2010). Dengan demikian, sudah jelas bahwa tujuan diadakannya pendekatan penilaian ini adalah untuk meningkatkan proses kemampuan siswa dengan cara mengomunikasikan kekurangan ataupun kelebihan pada saat melakukan proses pembelajaran. Mari kita lihat narasi di bawah untuk menggambarkan situasi pelaksanaan jenis pendekatan *assessment for learning* ini:

Guru sedang mengajar bagaimana menulis teks deskriptif yang baik. Pada saat pelajaran berlangsung, guru memberikan tugas kepada siswa untuk mencoba membuat teks deskriptif sederhana dengan mendeskripsikan artis idolanya. Murid-murid pun setuju dan langsung bergegas mengambil alat tulis dan langsung mengerjakan tugas dari gurunya. Setelah beberapa menit, tugas pun selesai dan dikumpulkan. Selang beberapa saat guru pun selesai mengecek tugas-tugas siswanya. Pada saat mengecek jawaban para siswa guru memberikan tulisan-tulisan kecil di bawah pekerjaan siswa yang berisi saran atau kritik yang membangun yang memang secara prosedur siswa masih belum memenuhi standar penulisan teks deskripsi yang baik. Dalam feedbacknya tersebut, guru tidak mencoba untuk menjustifikasi kalau memang apa yang ditulis siswanya salah, namun guru tersebut memberi pengarahan kepada siswa agar tulisannya lebih baik lagi. Dan kemudian setelah proses pengecekan berakhir, guru pun membagikan kertas tugas kepada siswa-siswanya untuk dipelajari lagi agar ke depannya dapat menulis teks deskripsi dengan baik dan benar.

Contoh kasus di atas menggambarkan situasi di kelas bagaimana pendekatan ini berlangsung. Dapat terlihat jelas bahwa guru menggunakan penilaian untuk suatu pembelajaran kepada siswanya. Dengan memberikan masukan dan kritik yang membangun, hal tersebut jelas dilakukan untuk meningkatkan kemampuan dan motivasi siswanya dalam belajar.

Ternyata cara tersebut terbukti memiliki level yang signifikan untuk meningkatkan mutu belajar. Menurut Fulcher (2010) ada tiga hal yang harus diperhatikan guru pada saat melaksanakan pendekatan jenis ini.

1. Dalam memberikan *feedback*, guru mencoba untuk membuat siswa paham tentang apa yang telah dipelajari; apa yang ditulis guru dalam *feedback*-nya harus bersifat deskriptif jangan *evaluative* (menjustifikasi), sehingga tidak menimbulkan kesan negatif (Rea-Dickins, 2006).

2. Siswa harus tahu aspek mana yang harus ditingkatkan dan guru dianjurkan untuk memberi saran bagaimana cara meningkatkannya.
3. Tentunya siswa membutuhkan waktu untuk menggali dan merespons *feedback* yang telah diberikan oleh guru. Guru dianjurkan untuk memberikan waktu kepada siswa agar mereka mengembangkan kemampuannya dengan caranya sendiri.

Ketiga hal ini memang benar-benar patut menjadi acuan bagaimana kita menjalankan pendekatan penilaian untuk pembelajaran. Seperti yang dapat kita lihat pada poin pertama, di mana jangan sampai ada kesan negatif dari siswa. Maksudnya, dalam proses penilaian siswa terkadang merasa tertekan apabila guru terus-menerus menyalahkan apa yang telah siswa kerjakan di dalam kelas apalagi jika siswa sampai mendapatkan intimidasi. Hal tersebut bukannya meningkatkan kemampuan siswa, malah siswa menjadi takut dan tidak mau untuk belajar apa yang sudah diajarkan guru. Tentu pada poin yang kedua dimaksudkan bahwa dengan membangun komunikasi yang baik, dengan cara memberi saran, kesan, dan kritik secara positif siswa akan termotivasi untuk meningkatkan pembelajarannya. Dalam poin ketiga dapat kita lihat bahwa guru harus benar-benar mengatur waktu untuk memberi ruang kepada siswa dalam merespons atau mempelajari materi sesuai apa yang telah disarankan oleh guru.

Biasanya dalam dua jam pelajaran belum tentu cukup untuk menerapkan ketiga poin di atas. Untuk menghadapi situasi seperti ini, guru harus memutar otaknya untuk mencari cara agar apa yang sudah menjadi tujuan pembelajarannya itu tercapai sesuai waktu yang ditentukan. *Collaborative learning* atau belajar secara berkolaborasi mungkin menjadi solusi yang efektif untuk dilakukan dalam pembelajaran di kelas, di mana siswa saling berdiskusi untuk memecahkan masalah dan fokus kepada apa yang mereka ingin tahu.

Biasanya selain menerapkan metode kolaborasi di atas, guru bisa menggunakan bantuan untuk melaksanakan penilaian jenis ini. Bantuan yang dimaksud berupa *record keeping* atau buku pencatatan *progress* siswa. Dalam bukunya, Lorna (2003) menjelaskan bahwa buku catatan progres tersebut terbukti sangat membantu untuk merekam kemampuan siswa dari waktu ke waktu. Contoh dari *record keeping* adalah portofolio siswa, artefak, atau lembar kerja siswa untuk mengikuti jejak dalam rangkaian pembelajarannya.

3. Penilaian Sebagai Pembelajaran (*Assessment as Learning*)

Pada pendekatan *assessment as learning* ini, intinya hampir sama seperti pendekatan sebelumnya di mana masih menggunakan bentuk *formative assessment* dalam proses pembelajarannya. Perbedaan pada pendekatan jenis ini adalah keterlibatan siswa dalam melakukan penilaian. Jadi, bukan guru lagi yang memberi penilaian terhadap siswa, namun siswa berhak menilai dirinya sendiri atau temannya. Pada pendekatan *assessment as learning* ini siswa juga dilibatkan dalam merumuskan prosedur penilaian, kriteria, maupun rubrik atau pedoman penilaian sehingga mereka mengetahui dengan pasti apa yang harus dilakukan agar memperoleh capaian belajar yang maksimal. Seperti yang dikatakan oleh Fulcher (2010) bahwa guru meminta siswa untuk membuat kriteria penilaiannya sendiri dalam satu grup, dan kemudian digunakan untuk menilai satu sama lain. Lorna (2003) juga mengungkapkan bahwa kegiatan tersebut merupakan proses dari metakognisi di mana siswa itu sendirilah yang memonitor apa yang sedang dipelajari, menggunakan *feedback*-nya sendiri untuk membuat suatu penyesuaian apakah yang ia pahami sudah sesuai dengan materi yang diajarkan. Tentu pendekatan penilaian jenis ini merupakan pendekatan di mana siswa merupakan asesor untuk dirinya sendiri.

Transparansi menjadi hal yang sangat penting dalam mengaplikasikan pendekatan jenis ini. Jika *assessment as learning* ini bisa berjalan dengan baik dan sukses, maka tingkat transparansi dalam penilaian pun tergolong tinggi. Sama seperti yang diutarakan Lorna sebelumnya, Fulcher (2010) juga menyatakan bahwa praktik dari pendekatan ini menggunakan kriteria yang transparan yang didesain untuk membantu siswa mengembangkan kesadarannya bahwa apa yang sedang dihasilkannya sekarang sudah sesuai atau belum dengan tujuan pembelajaran yang diberikan oleh guru. Proses tersebutlah yang dapat meningkatkan pengetahuan siswa dalam sebuah pembelajaran.

Untuk dapat melihat lebih jelas mengenai penerapan pendekatan penilaian ini, mari kita lihat contoh kasus di bawah ini.

Di suatu kelas yang berisi sekitar 20 siswa sedang belajar Bahasa Inggris dengan materi pronunciation. Sebelum guru memulai kegiatan belajarnya, guru meminta siswa untuk membuat grup terlebih dahulu. Siswa diberi arahan oleh guru untuk membuat sebuah panduan penilaian dengan mencakup beberapa aspek. Setelah

semuanya siap, guru meminta semua siswanya untuk membaca materi tentang pronunciation tersebut, tentu dengan batas waktu yang ditentukan. Setelah waktu habis, guru meminta satu per satu siswa di setiap kelompok untuk menyampaikan materi yang telah dibaca dan otomatis teman dalam satu kelompoknya menilai penampilan dari siswa tersebut dan memberi feedback positif kepadanya. Jadi, kegiatan ini semacam kegiatan diskusi namun ada proses penilaian yang transparan di dalamnya. Dan setelah semuanya selesai, guru meminta siswa untuk mengumpulkan hasil nilai yang telah mereka buat pada saat proses pembelajaran tadi. Di situlah peran guru untuk melihat progres dari siswanya.

Contoh kasus di atas menggambarkan situasi di kelas di mana proses pendekatan *assessment as learning* berlangsung. Terlihat di mana siswa saling memberi nilai kepada siswa lainnya dan guru hanya menginstruksikan apa yang harus siswa kerjakan. Terlihat bahwa proses kegiatan belajar mengajar seperti contoh di atas layaknya forum diskusi yang sedang berlangsung. Namun jika kita lihat lebih jauh ada proses penilaian di situ di mana siswa membuat catatan nilai yang sesuai dengan rubrik mereka dan juga mereka memberikan *feedback* kepada temannya.

Table 4.1 Pendekatan-pendekatan dalam Penilaian

Pendekatan	Tujuan	Perbandingan	Asesor Utama
<i>Assessment of learning</i>	Untuk memberikan sebuah keputusan; <i>ranking</i> siswa, promosi, dan lain-lain	Siswa lain	Guru
<i>Assessment for learning</i>	Sebagai informasi untuk keputusan instruksional guru	Kurikulum atau kompetensi yang harus dicapai	Guru
<i>Assessment as learning</i>	Sebagai pemonitor dan pengoreksi kemampuan diri sendiri	Tujuan dari siswa itu sendiri dan kurikulum atau kompetensi yang harus dicapai	Siswa

Tabel di atas merupakan ringkasan dari ketiga pendekatan yang digunakan dalam penilaian, yaitu *assessment of learning*, *assessment for learning*, dan *assessment as learning*. Ketiga pendekatan di atas memang sama-sama berkontribusi dalam mengembangkan dan meningkatkan pembelajaran siswa, namun melalui cara dan trik yang berbeda-beda. Tentu tidak ada yang salah apakah kita sebagai guru mau menggunakan

salah satu dari ketiga pendekatan tersebut atau bahkan menerapkan semua pendekatan tersebut. Terpenting dalam melakukan penilaian harus transparan dan adil atau objektif.

Latihan

Pikirkan tentang bagaimana Anda menguji atau menilai siswa Anda sendiri di kelas.

1. Buat daftar jenis tes apa yang Anda lakukan. Seberapa sering Anda melakukan penilaian pembelajaran (*Assessment of Learning*), penilaian untuk pembelajaran (*assessment for learning*), dan penilaian sebagai pembelajaran (*assessment as learning*)!
2. Diskusikan dengan rekan kerja Anda mengapa Anda lebih sering melakukan penilaian jenis tertentu!

Bab 5

Merancang Tes

Objek Pembelajaran

- A. Siklus Perancangan Tes
- B. Definisi Konstruksi Model
- C. Model Kompetensi Komunikatif
- D. Mendesain Spesifikasi Tes

Menentukan apa yang harus dinilai bukan perkara yang mudah. Bukan sekadar menentukan materi yang akan diujikan kepada siswa, namun guru juga harus benar-benar mengetahui tentang tujuan diadakan tes tersebut. Contohnya, mengapa mengadakan tes tersebut, menguji coba tes yang akan diberikan, dan sebagainya. Bachman & Palmer (1996) menegaskan bahwa ada tiga langkah dalam membuat dan menggunakan tes tersebut, yang pertama adalah perencanaan (desain), operasionalisasi, dan yang terakhir adalah administrasi (pelaksanaan). Maka dari itu, pada bab ini kita akan membahas tentang apa saja yang harus dites.

A. Siklus Perancangan Tes

Dalam tahap ini *test-makers* atau si pembuat tes harus benar-benar mengerti apa yang akan diujikan kepada siswa. Tentu ini sangat krusial, karena jika kita tidak mengetahui tujuan kita melakukan tes, maka format dan isi dari tes tersebut tidak akan bisa mengukur sebuah kemampuan. Fulcher (2010) juga menyatakan bahwa tanpa pernyataan yang jelas mengenai tujuan diadakannya tes, sulit bagi pengembang tes untuk menentukan isi dan format dari tes tersebut. Jadi, pada tahap ini harus benar-benar dipahami untuk apa, mengapa, dan untuk siapa tes diadakan. Untuk mempermudah menentukan tujuan dalam pengadaan sebuah penilaian dan tes, Fulcher (2010) mengutip beberapa pertanyaan dari Harris dan Bell (1994) sebagai acuannya, di antaranya:

1. Apa yang ingin kita capai?
2. Bagaimana permasalahan diidentifikasi?
3. Siapa yang seharusnya terlibat?
4. Apa yang mungkin menjadi efek dari evaluasi?
5. Bagaimana informasi didapatkan dan dianalisis?
6. Siapa yang seharusnya mengumpulkan dan menganalisis informasi?
7. Sumber informasi apa yang dapat digunakan?
8. Apa yang menjadi batasan (e.g. waktu, tenaga, regulasi)?

Beberapa pertanyaan tersebut lalu dikerucutkan agar menjadi tujuan dari pelaksanaan sebuah tes ataupun penilaian dan menjadi landasan untuk apa yang akan dilakukan ke depannya, seperti bagaimana cara mengolah skor, soalnya akan seperti apa, dan sebagainya.

Dalam menghubungkan apakah soal yang kita buat sudah sesuai dengan materi atau pembelajaran, Fulcher (2010) mempunyai beberapa cara. *Pertama*, dalam membuat soal kita dianjurkan mengarah kepada contoh-contoh di silabus dengan melihat proses dan keterampilan yang menjadi target pembelajarannya. *Kedua*, dengan melihat tujuan pembelajaran atau materi yang kita sampaikan.

Setelah tujuan tes sudah jelas, selanjutnya yaitu menentukan kriteria tes. Dalam tahap ini, pembuat tes atau guru dituntut untuk menggunakan tes yang tepat agar hasilnya nanti sesuai dengan objektifnya. Sebagai contoh, ketika guru ingin melihat kemampuan siswa dalam berbicara Bahasa Inggris dia bisa menggunakan tes

wawancara, atau mungkin bermain peran sebagai kriteria tes yang digunakan.

B. Definisi Konstruk Model

Kita berada pada tahap ketiga dalam mendesain sebuah tes. Menurut Fulcher (2010) tahap ini merupakan tahap tersulit dalam mendesain sebuah tes. Mengapa tahap ini dibilang sulit? Mari kita pelajari lebih lanjut mengenai definisi dari model. Dalam menentukan apa yang harus dinilai tentu kita harus mengetahui latar belakang pengetahuan dari siswa yang kita uji. Artinya, kita harus tahu cara berpikir siswa ketika mereka belajar Bahasa Inggris, khususnya. Bagaimana mereka mengidentifikasi suatu masalah, bagaimana mereka menyikapi masalah-masalah yang ada itu penting untuk kita ketahui sebagai seorang guru.

Tentu hal ini sangat berhubungan dengan istilah autentik di mana soal-soal yang kita buat untuk sebuah tes sesuai dengan *real life* dari siswa yang kita uji. Maksudnya, dalam membuat butir soal sebisa mungkin kita harus menghubungkan isi dari soal tersebut dengan konteks yang memang siswa familiar dan paling penting sesuai dengan cara berpikirnya.

Jadi, hubungan antara isi soal dengan fenomena yang dialami siswa akan menjadikan tes tersebut autentik sesuai dengan konteks yang dialami siswa dalam pembelajaran. Misalnya, ketika guru membuat tes membaca, guru harus mencari materi atau teks-teks yang memang porsinya sesuai dengan pengetahuan siswanya dan sesuai dengan *trend* yang sedang berlangsung agar siswa dapat dengan senang mengidentifikasi masalah yang ada di dalamnya. Jadi, keseluruhan proses bisa didukung dengan menggunakan latar belakang pengetahuan yang relevan dan familiar dengan teks yang dilibatkan dalam tes tersebut.

Contoh lain ketika guru sedang membuat sebuah tes Bahasa Inggris untuk siswa SMK jurusan perhotelan. Pada konteks tersebut, guru harus membuat soal seautentik mungkin agar siswa dengan mudah mengilustrasikan dengan latar belakangnya. Dengan cara memberikan bacaan tentang lowongan pekerjaan untuk perhotelan, atau mungkin dengan memberikan bagaimana cara menjadi resepsionis hotel yang sesuai prosedur, dan tentu semua hal yang berhubungan dengan perhotelan.

Menurut McNamara (1996), ada tiga faktor yang harus diperhatikan dalam model keterampilan berbahasa, di antaranya: (1) apa maksudnya mengetahui suatu bahasa (model pengetahuan); (2) faktor yang mendasari kemampuan menggunakan bahasa (model kinerja); (3) bagaimana kita tahu contoh detail dari penggunaan bahasa (penggunaan bahasa yang sebenarnya).

Pada poin pertama bisa disebut sebagai pengetahuan bahasa siswa dan yang kedua adalah kemampuan siswa untuk menggunakannya. Kedua poin tersebut merupakan faktor yang memengaruhi kemampuan kita untuk berkomunikasi dan biasa disebut sebagai kompetensi komunikatif atau kemampuan komunikatif bahasa. Ketiga faktor di atas juga dijadikan sebagai landasan untuk menghubungkan antara nilai yang diperoleh dengan kemampuan siswa. Kesimpulannya adalah penggunaan model yang tepat akan membuat tes menjadi efektif dan efisien. Selain itu, hal tersebut akan menunjukkan tentang validasi dari soal yang kita buat sebagai bentuk dari pengukuran seberapa validnya soal yang kita buat.

C. Model Kompetensi Komunikatif

Model ini merupakan model yang sangat berpengaruh karena bermanfaat dan efektif pada pengajaran bahasa kedua dan juga dapat membuat sebuah tes menjadi lebih valid dan andal.

Tahap awal saat mendesain sebuah tes, yaitu tujuan tes. Setelahnya tujuan kita masuk ke tahap kriteria apa yang kita akan gunakan dalam tes yang kita rancang. Setelah ditentukan kriteria tesnya, kita menentukan definisi konstruksinya; mempertimbangkan soal yang akan digunakan dengan apa yang siswa alami sehari-hari (otentik), baru setelah itu guru mendesain soal, mengevaluasi soal yang sudah dibuat dengan cara mengujikannya sebagai uji coba apakah soal sudah valid atau andal, kemudian diujikan di lapangan untuk perakitan format terakhir, menyimpulkan kalau memang itu sudah siap diteskan, dan yang terakhir keputusan untuk diujikan. Siklus ini terus berputar seiring guru membuat tes di kelasnya.

D. Mendesain Spesifikasi Tes

1. Apa Itu Kisi-kisi?

Seperti yang kita tahu bahwa spesifikasi tes yaitu suatu format yang berisi tentang kriteria-kriteria dari sebuah tes atau soal dalam tes. Kita biasa menyebut format ini sebagai kisi-kisi. Kisi-kisi adalah format yang paling detail dalam mendesain sebuah tes (Fulcher, 2010) atau dalam istilah lain kisi-kisi juga dapat disebut sebagai *blueprint* (Fulcher, 2010; Alderson, Clapham & Wall, 1995). Kisi-kisi menuntun kita untuk mengetahui bagaimana menyusun soal ujian, menyusun *lay out* tes, dan bagaimana menempatkan bagian-bagian soal yang sulit. Dalam membuat sebuah kisi-kisi tidak semudah apa yang kita bayangkan.

Menurut Fulcher & Davidson (2007), peran utama dari kisi-kisi yaitu untuk membuat kesetaraan tes. Maksudnya, ketika kita ingin membuat sebuah soal dengan materi yang sama, tingkat kesulitan yang sama, objektif pengujian yang sama, namun isinya berbeda. Tentu dalam membuat soal, kita menginginkan sebuah perubahan bentuk-bentuk soal dengan memvariasikan konten yang ada dalam soal tanpa mengubah hasil dari soal tersebut. Dalam hal ini, kisi-kisi mempunyai efek yang besar dalam mengembangkan sebuah tes (Davidson & Lynch, 2002). Pada bab ini, kita akan membahas konsep umum dalam mendesain kisi-kisi.

2. Perencanaan dalam Pembuatan Tes

Dalam langkah pertama ini guru harus tahu pasti untuk apa tes diadakan dan untuk mengukur apa dan siapa tes tersebut. Tentunya soal-soal yang dibuat nanti harus sesuai dengan kurikulum yang digunakan di sekolah. Waktu pelaksanaan tes juga harus disesuaikan dengan regulasi yang ada. Untuk di Indonesia biasanya soal disesuaikan dengan kompetensi dasar melalui beberapa indikator yang sudah ada. Sama halnya dengan membuat kisi-kisi. Kita atau guru juga harus melihat kompetensi dasar sebagai acuan untuk membuat kisi-kisi soal, agar ujian yang diberikan sesuai dengan target atau tujuan pembelajaran.

3. Petunjuk Atau Instruksi Tes

Langkah ini merupakan sebuah deskripsi tentang apa yang siswa harus lakukan atau apa yang siswa harus jawab dalam setiap soal. Selain itu,

dalam petunjuk kisi-kisi juga menjelaskan bagaimana jawaban siswa dinilai. Misalnya, pada penskoran soal pilihan ganda itu cukup mudah; salah poinnya 0, jika benar poinnya 1. Lain halnya jika kita dihadapkan dengan soal esai, tentu *rating scale*-nya akan lebih rumit dibanding soal pilihan ganda tadi.

4. Penyusunan Tes

Pada tahap ini berisi tentang aturan bagaimana seluruh soal dalam tes tersebut disusun (Fulcher, 2010). Maksudnya, kita harus tahu soal akan mengacu dari mana dan ada berapa soal dalam tes yang kita buat dalam 1 materi belajar. Misalnya, seperti kita tahu biasanya dalam soal ujian ada empat jenis soal namun kita harus mengidentifikasi kira-kira berapa dari setiap jenis soal yang kita butuhkan pada tes yang kita buat.

Tahap ini memiliki peran yang cukup krusial dalam menentukan jumlah soal dan mengukur tingkat akurasi dari setiap soal apakah sudah sesuai dengan kompetensi-kompetensi yang harus dinilai atau belum.

5. Penyajian

Tahap ini terkadang sering luput atau dilewatkan oleh para guru pada saat mendesain sebuah tes. Mengapa demikian? Tes yang baik tidak serta merta berisi tentang materi pembelajaran yang akurat dengan tujuan pembelajarannya, namun dalam konteks penyajiannya juga harus diperhatikan. Padahal seharusnya di dalam kisi-kisi sudah dijelaskan bagaimana menyajikan sebuah tes. Misalnya dalam penyajian tes menggunakan kertas (*paper-pencil test*), Fulcher (2010) memberi contoh bahwa kita harus menentukan ukuran *margin* kertas, jenis *font* yang dipakai, spasi, dan tentu saja halamannya. Selain itu, Beliau juga menjelaskan bahwa menggunakan *computer-based test* sangat penting untuk memperhatikan penyajian dalam tes. Hal yang perlu diperhatikan untuk dalam penyajiannya yaitu seperti desain *interface*-nya, warna-warna dalam tesnya, navigasi, *scrolling*, dan banyaknya teks yang tersedia dalam layar.

Penyajian bukan serta merta variasi atau hiasan belaka dalam pelaksanaan sebuah tes. Kejelasan tulisan, jenis kertas yang dipakai, bahkan tampilan layar pada komputer pun dapat memengaruhi nilai siswa (Fulcher, 2010). Maka dari itu, penyajian merupakan faktor

penting yang harus diperhatikan oleh guru ketika menyajikan sebuah tes kepada siswanya.

E. Pelaksanaan

Hal yang harus diperhatikan dalam tahap ini yaitu bagaimana guru mengatur keamanan dan alokasi waktu dalam sebuah tes. Misalnya dalam mengatur duduk siswa agar diregangangkan, mengatur jumlah pengawas tes, boleh tidaknya menggunakan kamus dalam tes Bahasa Inggris atau dalam tes komputer diterapkannya *security system* untuk menghindari dari kecurangan. Selain itu, pembagian waktu dalam setiap sub tes juga harus diperhatikan. Misalnya untuk pilihan ganda 20 menit, uraian 15 menit, dan esai 10 menit. Hal tersebut harus diperhatikan oleh para guru agar proses tes dapat berlangsung dengan aman dan sesuai waktu yang ditentukan.

Latihan

1. Apa yang harus dilakukan oleh seorang guru jika ada beberapa tujuan pembelajaran tapi tidak bisa dicakup seluruhnya dalam sebuah tes?
2. Perlukah guru memberitahu siswa kisi-kisi soal? Mengapa?

[Halaman ini sengaja dikosongkan]

Bab 6

Bentuk-bentuk Penilaian

Objek Pembelajaran

- A. Penilaian Respons Terpilih (*Selected Response*)
- B. Penilaian Respons Terstruktur (*Constructed Response*)
- C. Penilaian Respons Pribadi (*Personal Response*)

Pada bab ini, akan dipaparkan berbagai jenis penilaian terutama penilaian keterampilan berbahasa yang dapat menjadi pengetahuan dan contoh bagi guru dalam melaksanakan penilaian di kelas. Secara umum penilaian dikategorikan dalam tiga klasifikasi, yaitu (1) penilaian respons terpilih; (2) penilaian tanggapan terstruktur; dan (3) penilaian respons pribadi. Paparan tentang masing-masing kategori akan diikuti dengan penjelasan mengenai kelebihan dan kekurangan masing-masing jenis penilaian.

A. Penilaian Respons Terpilih (*Selected Response*)

Kategori penilaian ini mengharuskan siswa untuk memilih jawaban yang benar di antara serangkaian pilihan yang terbatas. Dalam penilaian respons terpilih, siswa biasanya tidak membuat atau menunjukkan keterampilan bahasa apa pun. Dengan demikian, penilaian ini paling tepat untuk mengukur keterampilan reseptif seperti mendengarkan dan membaca.

Secara umum, penilaian respons terpilih relatif cepat untuk dilakukan. Selain itu, penilaian kategori ini relatif cepat, mudah, dan relatif objektif. Namun, penilaian ini memiliki dua kelemahan: (a) cukup sulit untuk dibuat, dan (b) tidak mengharuskan siswa untuk menggunakan bahasa produktif apa pun. Contoh dari penilaian kategori ini adalah benar-salah, menjodohkan, dan pilihan ganda.

1. Benar Salah

Jenis penilaian benar-salah menyajikan sampel bahasa dan mengharuskan siswa untuk merespons bahasa itu dengan memilih salah satu dari dua pilihan, benar atau salah. Kelebihan utama dari penilaian benar-salah adalah mereka berfokus pada kemampuan siswa untuk memilih jawaban yang benar dari dua alternatif. Jadi, penilaian benar-salah memberikan indikasi sederhana dan langsung apakah permasalahan telah dipahami. Namun demikian, satu hal tentang penilaian benar-salah adalah bahwa untuk menghasilkan item yang mendiskriminasi dengan baik, pembuat soal mungkin tergoda untuk menulis item yang rumit, yaitu yang menghidupkan arti satu kata atau frasa atau yang ambigu. Sebagian besar guru lebih suka membuat penilaian langsung di mana siswa yang tahu jawabannya mendapatkannya benar dan siswa yang tidak tahu jawabannya salah. Permasalahan lain adalah faktor tebakan (*guessing*) yang relatif besar. Dalam hal ini, peserta ujian memiliki peluang 50% untuk menjawab dengan benar bahkan jika mereka tidak tahu jawabannya. Namun, jika sejumlah besar item benar-salah yang dirancang dengan cermat digunakan, skor keseluruhan harus mengatasi sebagian besar pengaruh faktor menebak.

2. Menjodohkan/Padanan

Penilaian yang sesuai menyajikan kepada siswa dua daftar kata atau frasa; dan dari daftar tersebut siswa harus memilih kata atau frasa dalam satu daftar yang cocok dengan yang ada di daftar lainnya. Kelebihan utama dari penilaian padanan adalah bahwa faktor tebakan yang rendah (misalnya, hanya 10% untuk 10 item jika opsi tambahan disediakan). Penilaian padanan umumnya dibatasi untuk mengukur kemampuan siswa untuk mengaitkan satu set fakta dengan yang lain, yang dalam pengujian bahasa biasanya berarti mengukur pengetahuan kosakata pasif (misalnya kemampuan siswa untuk mencocokkan definisi dengan item kosakata).

3. Pilihan Ganda

Penilaian pilihan ganda mengharuskan siswa untuk memeriksa pertanyaan atau pernyataan dan memilih jawaban yang paling tepat untuk menyelesaikan kalimat atau mengisi bagian kosong dalam kalimat dari antara tiga, empat, atau lima pilihan. Penilaian pilihan ganda, seperti halnya penilaian menjodohkan/padanan, memiliki kelebihan, yaitu faktor menebak yang relatif kecil. Sedangkan penilaian benar-salah memiliki faktor menebak 50%, penilaian pilihan ganda biasanya memiliki faktor menebak 33%, 25%, atau 20% tergantung pada apakah ada tiga, empat, atau lima opsi. Penilaian pilihan ganda juga memiliki kelebihan yaitu bermanfaat untuk mengukur berbagai jenis poin pembelajaran yang cukup luas.

Namun demikian, penilaian pilihan ganda sering dikritik karena tidak selalu autentik, yaitu tidak mencerminkan penggunaan bahasa dalam kehidupan nyata. Dengan demikian, sebaiknya menghindari penggunaan penilaian pilihan ganda (atau penilaian benar-salah atau menjodohkan, dalam hal ini) untuk mengukur keterampilan produktif seperti menulis dan berbicara.

Meskipun demikian, banyak aspek bahasa, khususnya keterampilan reseptif, dapat diuji menggunakan penilaian pilihan ganda. Menguji membaca, mendengarkan, pengetahuan tata bahasa, dan diskriminasi fonem dengan item pilihan ganda dapat memberikan informasi yang berguna tentang kemampuan atau pengetahuan siswa di bidang-bidang tersebut dengan cukup efisien. Sayangnya, karena keterampilan

membaca, mendengarkan, dan tata bahasa sering kali merupakan satu-satunya penilaian yang diukur pada tes kecakapan yang biasa digunakan; penilaian pilihan ganda sering kali terlalu sering digunakan. Dilihat secara terbalik, penggunaan item pilihan ganda (biasanya karena kemudahan administrasi dan penilaian serta objektivitas) mungkin sering membatasi jenis keterampilan bahasa yang diuji dalam membaca, mendengarkan, dan tata bahasa. Selain itu, item pilihan ganda kadang-kadang digunakan secara tidak tepat (misalnya, penilaian pilihan ganda kemampuan menulis).

B. Penilaian Respons Terstruktur (*Constructed Response*)

Penilaian tanggapan terstruktur mengharuskan siswa untuk menghasilkan bahasa dengan menulis, berbicara, atau melakukan sesuatu yang lain. Oleh karena itu, penilaian ini mungkin paling tepat untuk mengukur keterampilan berbicara dan menulis yang produktif. Penilaian tanggapan terstruktur juga dapat berguna untuk mengamati interaksi keterampilan reseptif dan produktif, misalnya interaksi mendengarkan dan berbicara dalam prosedur wawancara lisan atau interaksi membaca dan menulis dalam penilaian kinerja di mana siswa membaca dua artikel akademik dan menulis sebuah esai yang membandingkan dan membedakan mereka.

Ada pertukaran tertentu dalam memutuskan apakah akan menggunakan penilaian respons yang dipilih atau terstruktur. Sebagai contoh, item respons terpilih memungkinkan untuk menebak, tetapi mereka relatif objektif. Di sisi lain, item tanggapan terstruktur menghilangkan beberapa faktor menebak tetapi menciptakan masalah subjektivitas, terutama ketika penilaian melibatkan unsur manusia dalam memutuskan apa jawaban yang benar.

Faktor menebak tidak terlalu menjadi masalah pada penilaian tanggapan terstruktur. Walaupun demikian, penilain kategori ini tidak sepenuhnya bebas dari unsur menebak. Misalnya, pada ujian menulis, beberapa siswa mungkin mencoba menggunakan kata-kata kunci dalam *prompt* untuk menulis tentang di sekitar topik untuk menjawab dengan harapan mendapatkan sesuatu yang akan dihitung sebagai jawaban yang benar.

Secara umum, penilaian tanggapan terstruktur tidak memiliki faktor menebak, dan mengukur penggunaan bahasa produktif serta interaksi keterampilan reseptif dan produktif. Namun, penilaian mungkin relatif sulit dan memakan waktu. Penilaian tanggapan terstruktur juga mungkin cukup subjektif tergantung pada jenisnya. Tiga jenis penilaian tanggapan yang dibangun umumnya digunakan dalam pengujian: mengisi titik-titik, jawaban singkat, dan performansi/penilaian kinerja.

1. Melengkapi Kalimat

Penilaian isi menghadirkan pernyataan dengan bagian dari konteks dihapus dan diganti dengan titik-titik/kosong. Untuk menjawab, siswa diharuskan mengisi bagian yang kosong.

Penilaian ini memiliki kelebihan yaitu cukup mudah dibuat, fleksibel dalam hal penilaian, dan cepat dalam administrasi. Selain itu, seperti jenis tanggapan terstruktur lainnya, penilaian ini mengukur kemampuan siswa untuk benar-benar menghasilkan/produksi bahasa, meskipun sejumlah kecil bahasa, dan membuka kemungkinan menilai interaksi antara keterampilan reseptif dan produktif (misalnya, dalam sebuah tes menyimak, siswa harus mendengarkan suatu bagian sambil membacanya dan mengisi bagian yang kosong).

Adapun kekurangannya adalah fokusnya yang sempit yaitu terbatas pada pengujian satu kata atau frasa pendek. Masalah lain adalah bahwa titik-titik untuk diisi mungkin memiliki sejumlah jawaban yang mungkin.

2. Jawaban Singkat

Penilaian jawaban singkat mengharuskan siswa untuk meneliti pertanyaan atau pernyataan dan merespons dengan satu atau lebih frase atau kalimat. Kelebihan dari penilaian jawaban singkat adalah bahwa mereka mudah diproduksi dan relatif cepat untuk dilakukan. Salah satu kelemahan dari penilaian jawaban singkat adalah bahwa mereka fokus pada penilaian beberapa frase atau kalimat. Kekurangan kedua adalah bahwa jawaban bisa lebih dari satu, yang berarti bahwa jika *prompt* tidak dibuat dengan hati-hati, setiap siswa dapat menghasilkan jawaban yang sama sekali berbeda.

3. Performansi/Penilaian Kinerja

Penilaian kinerja mengharuskan siswa untuk mencapai perkiraan dari kehidupan nyata, tugas autentik, biasanya menggunakan keterampilan produktif berbicara atau menulis tetapi juga menggunakan keterampilan membaca atau menulis atau keduanya. Penilaian kinerja dapat dilakukan dalam berbagai bentuk, termasuk tugas yang cukup tradisional seperti penulisan esai atau wawancara atau tugas pemecahan masalah, permainan peran, dan diskusi kelompok. Singkatnya, menurut definisi, penilaian kinerja memiliki tiga persyaratan: (a) siswa diminta untuk melakukan semacam tugas, (b) tugas harus seautentik mungkin, dan (c) performansi biasanya dinilai oleh penilai yang berkualitas.

Kelebihan utama dari penilaian kinerja adalah bahwa siswa dapat menampilkan komunikasi autentik (setidaknya sejauh komunikasi autentik dapat diperoleh dalam situasi pengujian apa pun). Penilaian kinerja memberikan lebih banyak (a) ukuran kemampuan siswa untuk merespons tugas dunia nyata, (b) perkiraan kemampuan siswa yang sebenarnya daripada penilaian pilihan ganda standar yang terstandarisasi, dan (c) prediksi kinerja masa depan siswa dalam situasi kehidupan nyata. Penilaian kinerja juga dapat digunakan untuk menangkal efek negatif *washback* dari pengujian standar, seperti bias dan konten yang tidak relevan. Bahkan, penilaian kinerja yang dirancang dengan baik dapat memberikan efek *washback* positif yang kuat (lihat diskusi di bawah), terutama jika mereka terkait langsung dengan kurikulum tertentu.

Penilaian kinerja relatif sulit untuk dibuat dan relatif memakan waktu untuk dikelola. Biaya yang cukup besar juga dapat dikeluarkan dalam mengembangkan penilaian kinerja, mengelolanya, melatih penilai, pelaporan skor, dan sebagainya. Kelemahan lain adalah butuh waktu dan sumber daya untuk mengumpulkan dan menyimpan rekaman audio atau video pertunjukan, menyediakan peralatan khusus. Validitas mungkin menjadi masalah karena inkonsistensi penilai, jumlah pengamatan terbatas, subjektivitas dalam proses penilaian, dan sebagainya. Validitas juga dapat menjadi masalah karena (a) cakupan konten yang tidak memadai; (b) kurangnya generalisasi konstruk; (c) sensitivitas penilaian kinerja terhadap metode pengujian, jenis tugas, dan kriteria penilaian; (d) membangun representasi yang kurang, yaitu

masalah generalisasi dari beberapa pengamatan hingga seluruh spektrum pertunjukan kehidupan nyata; dan (e) varians konstruk-tidak relevan (misalnya karakteristik kinerja yang tidak ada hubungannya dengan kemampuan nyata siswa). Tes keamanan mungkin juga bermasalah karena sejumlah kecil *prompt* (masing-masing *prompt* mungkin sangat mudah bagi peserta ujian untuk mengingat dan meneruskan ke yang lain), kesulitan membuat dan menyamakan *prompt* baru untuk setiap administrasi, dan efek potensial dari pengajaran ke ujian.

C. Penilaian Respons Pribadi (*Personal Response*)

Seperti penilaian tanggapan terstruktur, penilaian tanggapan pribadi mengharuskan siswa untuk benar-benar menampilkan kemampuan mereka, tetapi penilaian tanggapan pribadi juga memungkinkan tanggapan masing-masing siswa sangat berbeda. Dalam arti sebenarnya, penilaian ini memungkinkan siswa untuk mengomunikasikan apa yang ingin mereka komunikasikan. Secara umum, penilaian respons pribadi bermanfaat karena memberikan penilaian pribadi atau individual, dapat secara langsung terkait dengan dan diintegrasikan ke dalam kurikulum, dan dapat menilai proses pembelajaran secara berkelanjutan selama masa pengajaran. Namun, penilaian tanggapan pribadi juga memiliki kelemahan umum karena relatif sulit untuk dilakukan dan diorganisir dan melibatkan penilaian subjektif. Jenis penilaian respons pribadi yang paling umum adalah konferensi, portofolio, dan penilaian diri sendiri dan sejawat.

1. Konferensi

Penilaian konferensi biasanya melibatkan siswa untuk bertemu guru, biasanya dengan janji temu, untuk membahas karya tertentu atau proses pembelajaran, atau keduanya. Lebih penting lagi, konferensi berbeda dari bentuk penilaian lain karena mereka fokus langsung pada proses dan strategi pembelajaran. Sebagai contoh, pertimbangkan serangkaian konferensi yang diadakan untuk membahas beberapa konsep tulisan siswa. Selama konferensi, fokusnya bisa pada pandangan siswa dan kekhawatiran tentang proses pembelajaran yang mereka alami saat memproduksi dan merevisi komposisi mereka.

Secara total, keuntungan dari konferensi adalah bahwa guru dapat menggunakannya untuk (a) mendorong refleksi siswa pada proses belajar mereka sendiri; (b) membantu siswa mengembangkan citra diri yang lebih baik; (c) mendatangkan pertunjukan bahasa pada tugas, keterampilan, atau poin bahasa lainnya; atau (d) menginformasikan, mengamati, membentuk, dan mengumpulkan informasi tentang siswa. Secara alami, kelebihan tersebut diimbangi oleh kerugian tertentu. Dalam kasus konferensi, kerugiannya adalah mereka relatif memakan waktu, sulit dan subjektif untuk tingkat, dan biasanya tidak mendapat nilai atau nilai sama sekali.

2. Portofolio

Penggunaan portofolio sudah umum digunakan dalam bidang fotografi dan desain grafis. Mereka mengumpulkan portofolio pekerjaan mereka untuk menunjukkan pekerjaan dan keterampilan mereka dalam bentuk yang ringkas dan nyaman. Portofolio pun kemudian dipakai dalam dunia pendidikan untuk mendorong siswa memilih, menyusun, dan menampilkan pekerjaan mereka. Definisi penilaian portofolio adalah koleksi yang disengaja dari setiap aspek pekerjaan siswa yang menceritakan kisah pencapaian, keterampilan, upaya, kemampuan, dan kontribusi mereka (Brown, 1994).

Literatur menyatakan setidaknya ada tiga keuntungan dalam melakukan penilaian portofolio, yaitu memperkuat pembelajaran siswa, meningkatkan peran guru, dan meningkatkan proses pengujian. Penilaian portofolio dapat memperkuat pembelajaran siswa karena mereka (a) memanfaatkan pekerjaan yang biasanya dilakukan di kelas; (b) fokus perhatian peserta didik pada proses pembelajaran; (c) memfasilitasi praktik dan proses revisi; (d) membantu memotivasi siswa, jika direncanakan dengan baik, karena mereka menyajikan serangkaian kegiatan yang bermakna dan menarik; (e) meningkatkan keterlibatan siswa dalam proses pembelajaran; (f) membina kolaborasi siswa-guru dan siswa-siswa; (g) menyediakan sarana untuk menetapkan standar minimum untuk pekerjaan di kelas dan kemajuan; dan (h) mendorong siswa untuk mempelajari bahasa logam yang diperlukan bagi siswa dan guru untuk berbicara tentang pertumbuhan bahasa.

Penilaian portofolio dapat meningkatkan peran guru ke tingkat yang mereka (a) memberikan guru dengan gambaran yang lebih jelas tentang pertumbuhan bahasa siswa, (b) mengubah peran guru (di mata siswa) dari yang musuh ke yang seorang pelatih, dan (c) memberikan wawasan tentang kemajuan setiap siswa.

Penilaian portofolio dapat meningkatkan proses pengujian sejauh mereka (a) meningkatkan keterlibatan siswa dan guru dalam penilaian; (b) memberikan kesempatan bagi guru untuk mengamati siswa menggunakan bahasa yang bermakna untuk menyelesaikan berbagai tugas autentik dalam berbagai konteks dan situasi; (c) mengizinkan penilaian berbagai dimensi pembelajaran bahasa (termasuk proses, tanggapan, dan kegiatan); (d) memberikan kesempatan bagi siswa dan guru untuk bekerja bersama dan merefleksikan apa artinya menilai pertumbuhan bahasa siswa; (e) menambah variasi informasi yang dikumpulkan siswa; dan (f) membuat cara guru menilai pekerjaan siswa lebih sistematis. Literatur juga membahas setidaknya lima kelemahan menggunakan penilaian portofolio: masalah keputusan desain, logistik, interpretasi, keandalan, dan validitas. Merancang masalah keputusan termasuk memutuskan (a) siapa yang akan menentukan kriteria penilaian, (b) bagaimana kriteria penilaian akan ditetapkan, (c) siapa yang akan menentukan apa yang akan mengandung portofolio, dan (d) berapa banyak kegiatan kelas autentik sehari-hari akan menjadi termasuk dalam portofolio. Masalah logistik mencakup menemukan (a) peningkatan waktu dan sumber daya yang diperlukan untuk mendukung penilaian portofolio, (b) cara untuk mengandalkan pelatihan dan kemampuan guru untuk mengimplementasikan penilaian portofolio, dan (c) waktu bagi guru untuk membaca dan menilai portofolio secara teratur sepanjang tahun sekolah sambil secara bersamaan membantu siswa mengembangkan portofolio itu. Masalah interpretasi meliputi (a) menilai prestasi siswa sebagaimana terwakili dalam portofolio mereka; (b) menetapkan standar dan menafsirkan portofolio dengan cara yang sama adilnya bagi semua siswa; (c) melatih para guru untuk membuat interpretasi yang adil; dan (d) melaporkan hasil penilaian portofolio sehingga semua pemirsa yang tertarik (misalnya siswa, orang tua, administrator, politisi) dapat memahaminya. Masalah validitas melibatkan (a) memastikan keandalan yang memadai di antara penilai dan kesempatan ketika peringkat terjadi; (b) mendorong

objektivitas; (c) mencegah kesalahan mekanis, terutama yang dapat memengaruhi keputusan; (d) standarisasi proses pemeringkatan dan pemeringkatan; dan (e) memastikan akses yang sama ke sumber daya untuk semua siswa. Masalah validitas meliputi (a) yang menunjukkan validitas portofolio untuk keperluan pengambilan keputusan tentang siswa; (b) menentukan seberapa memadai portofolio menggambarkan kerja, pengembangan, dan kemampuan siswa; (c) mengidentifikasi dan mengendalikan setiap variabel intervensi potensial yang dapat memengaruhi prestasi siswa; dan (d) memisahkan kemampuan siswa mana yang menyebabkan karakteristik kinerja berupa jumlahnya.

Walaupun demikian, guru harus memperhatikan jangan sampai tugas portofolio hanya menjadi kumpulan kertas tidak berguna. Portofolio dapat gagal jika tidak sesuai dengan tujuan pembelajaran, jika pedoman tidak diberikan kepada siswa, jika tinjauan berkala dan umpan balik sistematis tidak ada, dan sebagainya.

Pengembangan portofolio yang berhasil harus mengikuti sejumlah langkah seperti yang digambarkan oleh O'Malley dan Valdez Pierce (1996) berikut ini.

- a. Nyatakan tujuan dengan jelas. Guru harus menunjukkan bagaimana kegiatan portofolio terhubung, terintegrasi dengan tujuan kurikuler. Perhatikan kepada siswa bagaimana portofolio mereka akan mencakup bahan-bahan dari mata pelajaran dan bagaimana koleksi itu akan meningkatkan tujuan kurikuler.
- b. Berikan pedoman tentang bahan apa yang akan dimasukkan. Setelah tujuan telah ditentukan, sebutkan jenis aktivitas yang harus dimasukkan. Hal ini akan sangat membantu untuk memberikan arahan yang jelas tentang cara memulai karena banyak siswa tidak akan pernah menyusun portofolio dan mungkin bingung tentang apa yang harus dilakukan. Portofolio sampel dari siswa sebelumnya dapat membantu merangsang beberapa pemikiran tentang apa yang harus dimasukkan.
- c. Mengomunikasikan kriteria penilaian kepada siswa. Kriteria penilaian merupakan aspek penting dari pengembangan portofolio dan yang paling kompleks. Dua sumber—penilaian sendiri dan penilaian guru—harus dimasukkan agar siswa dapat menerima manfaat maksimal. Penilaian diri harus sejelas dan sesederhana mungkin.

- d. Tentukan waktu untuk pengembangan portofolio. Jika siswa merasa tergesa-gesa mengumpulkan materi dan merefleksikannya, efektivitas proses portofolio berkurang. Pastikan bahwa siswa memiliki waktu yang diluangkan untuk mengerjakan portofolio (termasuk waktu di dalam kelas) dan bahwa kesempatan Anda sendiri untuk konferensi tidak terganggu.
- e. Tetapkan jadwal berkala untuk *me-review* dan konferensi. Dengan melakukan itu, guru membantu siswa untuk tidak mengerjakan semuanya di akhir semester.
- f. Tentukan tempat yang dapat diakses untuk menyimpan portofolio. Sangat tidak nyaman bagi siswa untuk membawa koleksi karya tulis dan karya seni. Jika guru memiliki ruang kelas mandiri atau tempat di ruang baca atau perpustakaan untuk menyimpan materi, itu dapat memberikan pilihan yang baik.
- g. Memberikan *washback* positif-penilaian akhir. Ketika sebuah portofolio telah selesai dan akhir jangka waktu telah tiba, penjumlahan akhir dilakukan. Haruskah portofolio dinilai? diberikan skor numerik tertentu? Setiap kelebihan diimbangi dengan kekurangan. Misalnya, skor numerik berfungsi sebagai data yang nyaman untuk membandingkan kinerja di antara siswa. Namun demikian, untuk portofolio yang berisi karya tulis, direkomendasikan untuk menggunakan skala penilaian holistik mulai dari 1 hingga 6 berdasarkan kualitas. Skor semacam itu mungkin paling baik dilihat sebagai ekuivalen numerik dari nilai huruf.

Namun orang dapat mengkritik bahwa tidaklah tepat untuk memberikan nilai angka dan huruf terhadap hasil kreativitas siswa dalam portofolio. Evaluasi dianggap lebih tepat diterapkan. Evaluasi tersebut dapat mencakup penilaian akhir dari pekerjaan oleh siswa, dengan pertanyaan seperti yang tercantum di atas untuk penilaian sendiri suatu proyek, dan evaluasi naratif tentang kekuatan dan kelemahan yang dirasakan oleh guru. Evaluasi akhir tersebut harus menekankan kekuatan tetapi juga menunjukkan jalan menuju tantangan pembelajaran di masa depan. Jelas bahwa portofolio mendapatkan peringkat kepraktisan yang relatif rendah karena waktu yang dibutuhkan guru untuk merespons dan melakukan konferensi dengan siswa

mereka. Namun demikian, mengikuti pedoman yang disarankan di atas untuk menentukan kriteria untuk mengevaluasi portofolio dapat meningkatkan keandalan penilaian portofolio.

3. Jurnal

Bentuk lain dari penilaian tanggapan personal adalah jurnal. Jurnal adalah catatan dari pikiran, perasaan, reaksi, penilaian, ide, atau kemajuan seseorang menuju tujuan. Siswa dapat mengartikulasikan pemikiran mereka tanpa ancaman pemikiran tersebut dihakimi nanti (biasanya oleh guru). Jurnal dapat menjadi sarana pencapaian tujuan pedagogis penting: praktik dalam mekanisme penulisan, menggunakan tulisan sebagai proses “berpikir”, individualisasi, dan komunikasi dengan para guru. Pada saat yang sama, kualitas penilaian--penulisan jurnal telah mengambil peran penting dalam proses pembelajaran. Karena sebagian besar jurnal adalah--atau seharusnya--dialog antara siswa dan guru. Jurnal memberikan kesempatan unik bagi seorang guru untuk menawarkan berbagai jenis umpan balik. Di sisi lain ada kekhawatiran bahwa jurnal terlalu bebas untuk dinilai secara akurat. Dengan begitu banyak kemungkinan variabilitas, sulit untuk menetapkan kriteria untuk evaluasi. Menilai entri jurnal bukanlah ilmu pasti. Penting untuk mengubah kelebihan dan potensi kelemahan jurnal menjadi langkah umum dan pedoman positif untuk menggunakan jurnal sebagai instrumen penilaian. Misalnya:

- a. Memperkenalkan siswa pada konsep penulisan jurnal. Bagi banyak siswa, terutama mereka yang berasal dari sistem pendidikan yang mengecilkkan gagasan dialog dan kolaborasi guru-siswa, penulisan jurnal pada awalnya akan sulit. Dengan pemodelan, dan penetapan tujuan yang jelas, siswa dapat membuat transisi yang luar biasa ke dalam proses penulisan jurnal yang berpotensi. Siswa yang diperlihatkan contoh entri jurnal dan diberi topik dan jadwal penulisan tertentu akan merasa nyaman dengan prosesnya.
- b. Nyatakan tujuan jurnal. Integrasikan penulisan jurnal ke dalam tujuan kurikulum dengan beberapa cara, terutama jika entri jurnal menjadi topik diskusi kelas.

Penilaian Diri dan Teman Sebaya

Penilaian diri membutuhkan siswa untuk menilai bahasa mereka sendiri, baik melalui penilaian diri kinerja, penilaian diri pemahaman, atau penilaian diri observasi. Penilaian mandiri kinerja mengharuskan siswa untuk membaca suatu situasi dan memutuskan seberapa baik mereka dapat merespons hal tersebut. Demikian pula, penilaian mandiri tentang pemahaman mengharuskan siswa untuk membaca suatu situasi dan memutuskan seberapa baik mereka bisa memahaminya. Sebaliknya, penilaian diri observasi mengharuskan siswa untuk mendengarkan rekaman audio atau rekaman video dari kinerja bahasa mereka sendiri (mungkin direkam dalam situasi alami atau dalam permainan peran) dan memutuskan seberapa baik menurut mereka kinerja mereka. Varian penilaian diri adalah penilaian teman sebaya, yang mirip dengan penilaian mandiri kecuali bahwa, sebagaimana tersirat oleh label, siswa menilai bahasa rekan-rekan mereka.

Penilaian diri memiliki sejumlah keunggulan. *Pertama*, penilaian diri dapat dirancang untuk dikelola relatif cepat. *Kedua*, mereka pasti melibatkan siswa secara langsung dalam proses penilaian. *Ketiga*, pada gilirannya, keterlibatan tersebut dapat membantu siswa memahami apa artinya belajar bahasa secara mandiri. Akhirnya, baik keterlibatan siswa dan otonomi mereka yang lebih besar dapat secara substansial meningkatkan motivasi mereka untuk belajar bahasa tersebut.

Penilaian diri juga memiliki sejumlah kelemahan. Keakuratan penilaian diri sebagian besar siswa sering bervariasi tergantung pada keterampilan linguistik dan bahan yang terlibat dalam evaluasi. Nilai yang dinilai sendiri sering dipengaruhi oleh kesalahan subjektif karena catatan akademik masa lalu, aspirasi karier, harapan teman sebaya atau orang tua, kurangnya pelatihan dalam belajar mandiri, dan lain-lain. Kesalahan subjektif semacam itu mungkin dapat diatasi sampai batas tertentu jika rubrik penilaian yang digunakan siswa untuk menilai diri mereka menggambarkan situasi linguistik yang jelas dan konkret di mana mereka harus mempertimbangkan kinerja mereka dalam hal perilaku yang dijelaskan secara tepat. Namun, kesalahan subjektif semacam itu mungkin sulit diatasi dalam beberapa situasi, yaitu ketika konsekuensi dari penilaian diri menjadi bagian integral dari penilaian itu sendiri.

Latihan

1. Sebutkan bentuk-bentuk penilaian yang biasa Anda lakukan di kelas dan beri alasan mengapa Anda memilih bentuk penilaian tersebut!
2. Analisislah kekurangan dan kelebihan masing-masing bentuk penilaian yang Anda gunakan di kelas dan kemukakan tantangan dalam menerapkannya!

Bab 7

Penilaian Alternatif

Objek Pembelajaran

- A. Apa Itu Penilaian Alternatif?
- B. Mengapa Kita Harus Menggunakan Penilaian Alternatif?
- C. Kelebihan dan Kekurangan dari Penilaian Alternatif
- D. Pedoman untuk Membangun Penilaian Alternatif

Sebuah pertanyaan yang umum menghampiri kita sebagai guru dan menjadi dasar dari penilaian: “Bagaimana kita tahu bahwa siswa kita mempelajari apa yang seharusnya mereka pelajari?”. Yang dilakukan secara umum adalah dengan melakukan tes atau asesmen terhadap siswa melalui tes tradisional yaitu pensil dan kertas (*pencil and paper test*). Digunakan sebagai alat untuk mengaudit kemajuan siswa kami. Jenis tes ini biasanya tidak interaktif. Para siswa tidak diperbolehkan untuk melihat tes sebelum mengambilnya dan mereka tidak diizinkan untuk melihat bagaimana mereka akan dievaluasi terlebih dahulu.

A. Apa Itu Penilaian Alternatif?

Penilaian alternatif bukanlah tes kertas dan pensil. Sebaliknya, ini adalah proses interaktif antara guru dan siswa. Tes atau asesmen yang diberikan berupa aktivitas yang didasarkan pada pengalaman kehidupan nyata. Terpenting, siswa diberi kriteria penilaian terlebih dahulu sehingga mereka tahu persis bagaimana mereka akan dinilai.

B. Mengapa Kita Harus Menggunakan Penilaian Alternatif?

Banyak guru akan setuju bahwa tujuan pembelajaran adalah untuk mengajarkan siswa bagaimana menggunakan bahasa dalam situasi kehidupan nyata. Jika tujuannya adalah untuk mengembangkan keterampilan bahasa komunikatif, maka guru harus dapat menilai kinerja siswa menggunakan keterampilan tersebut. Hal ini dapat dicapai dengan menerapkan penilaian alternatif.

Para peneliti telah mengembangkan kriteria yang jelas untuk menciptakan dan menerapkan berbagai bentuk penilaian alternatif. Mereka percaya bahwa penilaian yang dikembangkan dengan baik mencakup membaca, menulis, berbicara, dan mendengarkan dalam bahasa target. Menggunakan penilaian alternatif di kelas bahasa membutuhkan tugas autentik, rubrik penilaian, dan banyak umpan balik. Tugas autentik mencerminkan situasi kehidupan nyata yang menarik bagi semua jenis gaya belajar. Dengan rubrik di tangan, siswa tahu apa yang perlu mereka lakukan untuk mencapai kesuksesan.

Dengan kata lain, penilaian alternatif adalah bentuk penilaian yang berbeda dengan tes yang lebih tradisional. Filosofi umum di baliknya adalah memberi peserta kesempatan untuk menunjukkan apa yang bisa mereka lakukan daripada menyoroti kekurangan mereka. Penilaian biasanya disusun berdasarkan tugas-tugas autentik yang memberikan kesempatan kepada peserta didik untuk menunjukkan kemampuan mereka. Misalnya, dalam kelas berbicara (*speaking*) peserta didik diberi kesempatan untuk fokus pada komunikasi daripada jawaban yang benar atau salah. Di kelas-kelas ini, peserta didik juga memiliki kesempatan untuk menilai diri mereka sendiri (*self-assessment*) dan teman sebaya (*peer-assessment*) mereka.

Penilaian alternatif digunakan untuk menentukan apa yang dapat dan tidak dapat dilakukan siswa, berbeda dengan apa yang mereka

lakukan atau tidak ketahui. Penilaian alternatif, juga disebut sebagai tes kinerja atau penilaian autentik, digunakan untuk menentukan apa yang dapat dan tidak dapat dilakukan siswa, berbeda dengan apa yang mereka lakukan atau tidak ketahui. Pengukuran penilaian alternatif menerapkan kecakapan lebih dari mengukur pengetahuan. Contoh tipikal penilaian alternatif meliputi portofolio, pekerjaan proyek, dan kegiatan lain yang memerlukan beberapa jenis rubrik.

Inti dari penilaian kinerja adalah siswa diberi kesempatan untuk melakukan satu atau lebih hal berikut ini.

1. Tunjukkan kemampuan mereka.
2. Lakukan tugas yang bermakna.
3. Menerima umpan balik oleh orang yang berkualifikasi dalam hal kriteria yang relevan dan dapat dipertahankan.

Singkatnya, tujuan untuk menggunakan penilaian alternatif adalah untuk menilai kemampuan siswa dalam melakukan tugas-tugas kompleks yang secara langsung terkait dengan hasil belajar.

C. Kelebihan dan Kekurangan dari Penilaian Alternatif

Penilaian alternatif menyediakan sarana untuk menilai keterampilan yang tidak dapat dinilai secara langsung dengan tes tradisional. Penilaian ini berfokus pada kinerja siswa dan kualitas pekerjaan yang dilakukan oleh siswa. Namun demikian, proses penilaian alternatif memakan waktu, upaya, peralatan, bahan, fasilitas, atau dana. Lebih lanjut lagi, proses penilaian terkadang lebih subjektif daripada ujian tradisional.

D. Pedoman untuk Membangun Penilaian Alternatif

Tetapkan tujuan instruksional yang ingin Anda nilai se jelas mungkin dalam hal isi materi pelajaran dan serangkaian keterampilan atau operasi yang akan ditunjukkan oleh seorang siswa yang terampil.

Bedakan antara hasil-hasil yang dapat dinilai secara valid hanya dengan penilaian kinerja dan yang dapat dinilai sama efektifnya dengan langkah-langkah objektif.

Siswa akan kesulitan menunjukkan langkah-langkah menari di atas kertas.

Buat tugas yang mendatangkan bukti kemampuan siswa untuk melakukan keterampilan yang ditargetkan. Selanjutnya, putuskan jenis bimbingan guru apa yang dapat digunakan sambil tetap memberi siswa kebebasan untuk belajar dan melakukannya dengan cara mereka sendiri. Siswa dapat melakukan langkah-langkah dalam urutan apa pun yang mereka inginkan. Guru dapat meletakkan nama-nama langkah yang berbeda di papan tulis untuk membantu siswa mengingatnya jika diperlukan. Cobalah penilaian dan buat revisi seperlunya. Revisi dapat mencakup memberikan instruksi dan harapan yang lebih rinci kepada siswa.

Latihan

1. Jelaskan perbedaan antara penilaian tradisional dan alternatif dan berikan contohnya! Dari contoh tersebut, identifikasi dan buatlah rentang penilaian, mulai dari penilaian yang sangat tradisional ke arah penilaian alternatif.
2. Jelaskan pendapat Anda tentang penilaian alternatif dalam hubungannya dengan prinsip penilaian yang autentik dan praktis!

Bab 8

Merancang dan Menggunakan Rubrik

Objek Pembelajaran

- A. Pengertian Rubrik
- B. Manfaat Rubrik
- C. Jenis-jenis Rubrik
- D. Komponen-komponen Rubrik
- E. Langkah-langkah Merancang Rubrik
- F. *Me-review* Rubrik
- G. Menggunakan Rubrik

A. Pengertian Rubrik

Rubrik adalah alat bagi guru untuk menetapkan kriteria penilaian untuk tugas. Tidak hanya alat ini berguna bagi para guru, alat ini juga bermanfaat bagi siswa. Rubrik mendefinisikan secara tertulis apa yang diharapkan dari siswa untuk mendapatkan nilai tertentu pada suatu tugas.

Andrade & Heritage (2018) mendefinisikan rubrik sebagai alat penilaian yang mencantumkan kriteria untuk sebuah karya atau hasil kerja siswa. Sebagai contoh, sebuah rubrik untuk esai dapat memberi tahu siswa bahwa pekerjaan mereka akan dinilai berdasarkan tujuan, organisasi, perincian, suara, dan mekanik.

Definisi lain dikemukakan oleh Brookhart (2013). Menurut Brookhart, rubrik adalah seperangkat kriteria yang koheren untuk menilai pekerjaan siswa yang mencakup deskripsi tingkat kualitas kinerja pada kriteria tersebut. Kedengarannya cukup sederhana, bukan? Sayangnya, definisi rubrik ini jarang ditunjukkan dalam praktik penilaian guru sehari-hari. Jika kita berselancar di internet, kita bisa menemukan banyak rubrik yang ditawarkan. Namun demikian, tidak sedikit rubrik yang tidak menggambarkan apa yang harus diukur atau dinilai. Patut diingat adalah bahwa rubrik adalah rubrik harus deskriptif dan tidak evaluatif. Tentu saja, rubrik dapat digunakan untuk mengevaluasi, tetapi prinsip operasinya adalah guru mencocokkan kinerja atau *performance* siswa dengan deskripsi yang sudah ditetapkan dalam rubrik daripada *men-judge* atau menghakimi tanpa kriteria yang jelas. Rubrik yang efektif memiliki kriteria yang sesuai dan deskripsi kinerja yang ditulis dengan baik.

B. Manfaat Rubrik

Tujuan utama rubrik adalah untuk menilai kinerja. Misalnya, guru mengamati siswa dalam proses melakukan sesuatu, seperti saat siswa menggunakan obeng listrik atau siswa sedang berpidato. Kinerja lain contohnya adalah guru mengamati produk yang merupakan hasil karya siswa, seperti pot keramik buatan siswa atau laporan tertulis.

Pada dasarnya semua kinerja siswa dapat dinilai dengan menggunakan rubrik. Dalam setiap penilaian, dengan panduan rubrik, guru mengamati seberapa tepat, seberapa lengkap, atau seberapa baik pertanyaan dijawab oleh siswa, misalnya. Namun demikian, ada satu hal yang tidak berfungsi dengan baik dengan rubrik yaitu pertanyaan dengan jawaban benar atau salah. Rubrik memberi struktur pada pengamatan. Menyesuaikan pengamatan Anda dengan karya siswa dengan berpedomankan deskripsi di rubrik. Hal ini mencegah guru terburu-buru untuk menghakimi yang dapat terjadi dalam situasi

penilaian kelas. Alih-alih menghakimi kinerja, rubrik menggambarkan kinerja. Penilaian terhadap kualitas kinerja siswa yang berdasarkan pada rubrik dapat digunakan untuk umpan balik dan pengajaran. Ini berbeda dari penilaian kualitas dari skor atau nilai yang diterima tanpa rubrik. Penghakiman tanpa deskripsi tidak dapat meningkatkan kualitas pembelajaran di ruang kelas.

Rubrik sangat penting karena rubrik menggambarkan kepada siswa kualitas dari tugas yang mereka kerjakan. Dengan demikian, rubrik membantu guru mengajar, membantu mengoordinasikan instruksi dan penilaian, dan membantu siswa belajar.

1. Rubrik Membantu Guru Mengajar

Untuk menulis atau memilih rubrik, guru perlu fokus pada kriteria penilaian pembelajaran. Fokus pada apa yang guru ingin siswa pelajari daripada apa yang ingin guru ajarkan. Jika hal ini dilakukan, rubrik dapat membantu meningkatkan pengajaran. Tanpa kejelasan apa yang siswa harus pelajari, sulit untuk mengetahui berapa banyak dari aspek konten yang akan diajarkan. Rubrik membantu kejelasan konten dan hasil pembelajaran.

Hal yang tak kalah pentingnya adalah rubrik mengurangi jumlah waktu yang dihabiskan guru untuk mengevaluasi pekerjaan siswa. Misalnya, saat siswa diberikan rubrik untuk menilai hasil karya atau tugas mereka sendiri atau menilai tugas temannya, siswa lebih mandiri untuk melihat kekuatan dan kelemahan mereka sendiri. Jadi, rubrik memberi siswa umpan balik yang lebih informatif tentang kekuatan dan bidang yang perlu ditingkatkan.

2. Rubrik Membantu Mengoordinasikan Pengajaran dan Penilaian

Sebagian besar rubrik harus dirancang untuk penggunaan berulang. Siswa diberi rubrik di awal pengajaran. Mereka melakukan tugas atau pekerjaan mereka, menerima umpan balik, berlatih, merevisi atau melakukan tugas lain, terus berlatih, dan akhirnya menerima nilai—semuanya menggunakan rubrik yang sama, yang memuat deskripsi tentang kriteria dan tingkat kualitas yang akan menunjukkan hasil pembelajaran.

3. Rubrik Membantu Belajar Siswa

Kriteria dan deskripsi tingkat kinerja dalam rubrik membantu siswa memahami seperti apa kinerja yang diinginkan dan seperti apa penampilannya. Rubrik yang efektif menunjukkan kepada siswa bagaimana mereka akan tahu sejauh mana kinerja mereka memenuhi target yang diharapkan. Dan jika digunakan secara formal juga dapat menunjukkan kepada siswa apa langkah mereka selanjutnya untuk meningkatkan kualitas kinerja mereka.

Rubrik dapat meningkatkan dan memantau siswa, dengan membuat harapan guru menjadi jelas dan dengan menunjukkan kepada siswa bagaimana memenuhi harapan tersebut. Hasilnya sering ditandai dengan peningkatan kualitas kerja siswa dan pembelajaran. Dengan demikian, argumen paling umum untuk menggunakan rubrik adalah mereka membantu mendefinisikan apa itu kualitas.

Alasan kedua bahwa rubrik berguna adalah bahwa rubrik membantu siswa menjadi hakim yang lebih bijaksana atas kualitas karya mereka sendiri dan orang lain. Ketika rubrik digunakan untuk memandu penilaian sendiri dan rekan, siswa menjadi semakin mampu mengenali dan memecahkan masalah dalam pekerjaan mereka sendiri dan satu sama lain. Latihan yang berulang dengan penilaian teman sebaya, dan terutama penilaian diri sendiri, meningkatkan rasa tanggung jawab siswa atas pekerjaan mereka sendiri.

C. Jenis-jenis Rubrik

Rubrik biasanya dikategorikan oleh dua aspek komposisi yang berbeda. Salah satunya adalah apakah rubrik memperlakukan kriteria satu per satu (analitik) atau secara keseluruhan (*holistic*). Lainnya adalah apakah rubrik itu bersifat umum dan dapat digunakan untuk beberapa tugas yang mirip atau khusus untuk tugas yang spesifik dan hanya berlaku untuk satu penilaian. Ada empat jenis rubrik yang umum dikenal.

1. Rubrik Analitik

Rubrik analitik menggambarkan kinerja siswa pada setiap kriteria secara terpisah. Untuk sebagian besar keperluan penilaian kelas, rubrik analitik adalah yang terbaik. Berfokus pada satu per satu kriteria lebih baik untuk pengajaran dan lebih baik untuk penilaian formatif karena siswa

dapat melihat aspek kinerja mereka; apa yang sudah baik dan apa yang masih membutuhkan perhatian dan peningkatan. Penggunaan rubrik analitik ini juga baik digunakan untuk penilaian sumatif yang juga akan digunakan untuk membuat keputusan tentang masa depan—misalnya, keputusan tentang cara menindaklanjuti pokok bahasan tertentu atau keputusan tentang strategi mengajar ke depannya. Untuk sebagian besar keperluan kelas, guru direkomendasikan untuk menggunakan rubrik analitik.

Jenis Rubrik

Jenis Rubrik	Definisi	Keunggulan	Kelemahan
Analitik	Setiap kriteria dinilai secara terpisah	<ol style="list-style-type: none"> 1. Memberikan informasi diagnostik kepada guru 2. Memberikan umpan balik formatif kepada siswa 3. Lebih mudah dihubungkan dengan topik yang diajarkan 4. Baik untuk penilaian formatif dan bisa diadaptasi untuk penilaian sumatif 	<ol style="list-style-type: none"> 1. Memakan lebih banyak waktu untuk pemberian skor 2. Memerlukan banyak waktu untuk mencapai reliabilitas antarpemilai
Holistik	Semua kriteria dinilai secara keseluruhan/ simultan	<ol style="list-style-type: none"> 1. Pemberian skor lebih cepat dibanding rubrik analitik 2. Tidak memerlukan banyak waktu untuk menghasilkan reliabilitas antarpemilai 3. Cocok untuk penilaian sumatif 	<ol style="list-style-type: none"> 1. Memakan lebih banyak waktu untuk pemberian skor 2. Memerlukan banyak waktu untuk mencapai reliabilitas antarpemilai

2. Rubrik Holistik

Rubrik holistik menggambarkan kinerja atau performa siswa dengan menerapkan semua kriteria pada saat yang sama dan memungkinkan penilaian menyeluruh tentang kualitas kinerja atau tugas siswa. Satu situasi di mana penggunaan rubrik holistik lebih baik daripada rubrik analitik adalah di mana siswa hanya menerima nilai dan tidak akan mendapatkan umpan balik lebih lanjut dari hasil penilaian sumatif akhir.

Beberapa ujian akhir semester atau sekolah termasuk dalam kategori ini. Penilaian dengan rubrik lebih cepat ketika hanya ada satu keputusan untuk dibuat, daripada keputusan terpisah untuk setiap kriteria.

3. Rubrik Umum

Rubrik umum menggunakan kriteria dan deskripsi kinerja yang digeneralisasi yang dapat digunakan untuk menilai beberapa tugas yang berbeda. Semua tugas yang dikerjakan mahasiswa harus merupakan contoh dari hasil belajar yang sama—misalnya, menulis atau memecahkan masalah. Kriteria menunjuk pada aspek hasil pembelajaran dan bukan pada komponen dari satu tugas tertentu. Misalnya, kriteria memuat hal umum seperti karakteristik pemecahan masalah dan bukan kriteria khusus pemecahan masalah tertentu (yang spesifik untuk tiap topik tertentu). Deskripsi kinerja atau performa adalah umum, sehingga siswa mempelajari kualitas umum dan tidak terisolasi, komponen khusus tugas.

Apa alasan menggunakan rubrik umum? Rubrik umum memiliki beberapa keunggulan dibandingkan rubrik khusus tugas.

- a. Dapat dibagikan kepada siswa di awal tugas, untuk membantu mereka merencanakan dan memantau pekerjaan mereka sendiri. Rubrik umum tidak “Berikan jawaban” untuk pertanyaan yang diberikan guru. Rubrik ini tidak mengandung informasi apa pun yang ‘membocorkan’ jawaban dari pertanyaan. Rubrik ini justru berisi deskripsi umum misalnya tentang kemampuan siswa menjelaskan sesuatu yang didukung dengan contoh dan bukti. Rubrik ini membantu siswa cara dan bukan jawaban dalam menyelesaikan tugas. Oleh karena itu, dari waktu ke waktu rubrik umum membantu siswa membangun konsep tentang apa artinya melakukan keterampilan dengan baik (misalnya, penyelesaian masalah yang efektif membutuhkan alasan yang jelas yang dapat dijelaskan dan didukung). Rubrik umum ini dapat digunakan dengan berbagai tugas. Karena rubrik umum memfokuskan siswa pada pengetahuan dan keterampilan yang mereka pelajari dan bukan pada tugas khusus yang mereka selesaikan.
- b. Dapat digunakan dengan banyak tugas yang berbeda, memfokuskan siswa pada pengetahuan dan keterampilan yang mereka kembangkan

dari waktu ke waktu. Karena rubrik umum memusatkan siswa pada pengetahuan dan keterampilan yang harus mereka peroleh, rubrik ini dapat untuk tugas apa pun yang menjadi bagian dari keseluruhan domain pembelajaran untuk hasil pembelajaran tersebut. Namun, serangkaian tugas ini menunjukkan pengetahuan dan keterampilan yang penting, dan mereka berkembang seiring waktu dan dengan latihan. Penulisan esai, pemecahan masalah, desain eksperimental, dan analisis sistem politik masing-masing keterampilan penting dalam disiplin ilmu masing-masing. Jika rubriknya sama setiap kali seorang siswa melakukan jenis pekerjaan yang sama, siswa akan mempelajari kualitas umum penulisan esai yang baik, pemecahan masalah, dan sebagainya. Jika rubrik berbeda setiap kali siswa melakukan jenis pekerjaan yang sama, siswa tidak akan memiliki kesempatan untuk melihat melampaui esai atau masalah tertentu. Pendekatan umum mendorong siswa untuk berpikir tentang membangun pengetahuan dan keterampilan umum daripada berpikir tentang pembelajaran sekolah dalam hal menyelesaikan tugas individu.

- c. Jelaskan kinerja siswa dalam istilah yang memungkinkan banyak jalan berbeda menuju sukses.
- d. Membuat guru lebih fokus pada pengembangan keterampilan belajar siswa alih-alih penyelesaian tugas.
- e. Tidak perlu ditulis ulang untuk setiap tugas, dengan kata lain, rubrik umum dapat digunakan berulang-ulang.

4. Rubrik Khusus

Rubrik khusus, sesuai namanya adalah rubrik yang spesifik untuk tugas kinerja yang digunakan. Rubrik khusus berisi jawaban atas suatu masalah, atau daftar fakta dan konsep yang seharusnya disebutkan oleh siswa. Apa alasan menggunakan rubrik khusus? Rubrik khusus berfungsi sebagai “arahan penilaian” bagi guru atau siapa saja yang menilai pekerjaan siswa. Karena rubrik khusus ini merinci unsur-unsur yang dicari dalam jawaban siswa untuk tugas tertentu. Menilai respons siswa dengan rubrik khusus ini termasuk pekerjaan dengan inferensi yang lebih rendah daripada menilai respons siswa dengan rubrik umum. Inferensi yang lebih rendah artinya bahwa guru atau penilai yang berbeda untuk memahami apa yang diinginkan dari tugas siswa. Karena alasan

ini, tidak dibutuhkan waktu yang lama untuk melatih penilai untuk mencapai tingkat keandalan penilaian yang dapat diterima menggunakan rubrik khusus ini untuk penilaian skala besar. Demikian pula, lebih mudah bagi guru untuk menerapkan rubrik spesifik secara konsisten dengan praktik minimal. Rubrik umum membutuhkan waktu lebih lama untuk belajar menerapkannya dengan baik. Namun, rubrik khusus hanya berguna untuk penilaian. Guru tidak dapat memberikan rubrik ini kepada siswa diawal pembelajaran karena ini berarti memberikan jawaban atas tugas yang diberikan. Oleh karena itu, rubrik khusus tidak berguna untuk penilaian formatif. Dengan demikian, gunakan rubrik khusus untuk penilaian tertentu yang khusus saja.

D. Komponen-komponen Rubrik

Secara umum, komponen-komponen sebuah rubrik terdiri dari kriteria dan rentang level/dimensi kualitas pencapaian dari kriteria yang sudah ditetapkan seperti terlihat pada tabel di bawah ini. Pada sisi kiri tabel rubrik adalah kriteria menggambarkan aspek-aspek kunci dari karya siswa atau produk. Pada bagian atas tabel rubrik, rentang level yang mengidentifikasi tingkat kinerja/kriteria. Di bawah setiap bagian skala penilaian, adalah indikator atau deskriptor konkret untuk setiap tingkat kriteria.

Komponen Sebuah Rubrik

Rentang Level/Dimensi Kualitas Pencapaian

	Level 1	Level 2	Level 3
Kriteria 1			
Kriteria 2			
Kriteria 3			
Kriteria 4			

Kriteria

Komponen Rubrik

1. Kriteria	2. Level/Dimensi			
	<i>Very Creative</i>	<i>Creative</i>	<i>Ordinary/ Routine</i>	<i>Imitative</i>
<i>Depth & Quality of Ideas</i>	<i>Ideas represent a startling variety of important contexts or disciplines</i>	<i>Ideas represent important concepts from different contexts or disciplines</i>	<i>Ideas represent important concepts from the same or similar contexts or disciplines</i>	<i>Ideas do not represent important concepts</i>
<i>Variety of Sources</i>	<i>Created product draws on a wide-ranging variety of sources, including different texts, media, resource persons, and/or personal experiences</i>	<i>Created product draws on a variety of sources, including different texts, media, resource persons, and/or personal experiences</i>	<i>Created product draws on a limited set of sources and media</i>	<i>Created product draws on only one source, and/or sources are not trustworthy or appropriate</i>
<i>Organization & Combination of Ideas</i>	<i>Ideas are combined in original and surprising ways to solve a problem, address an issue, or make something new</i>	<i>Ideas are combined in original ways to solve a problem, address an issue, or make something new</i>	<i>Ideas are combined in ways that are derived from the thinking of others (for example, of the authors in sources consulted)</i>	<i>Ideas are copied and restated from the source(s) consulted</i>
<i>Originality of Contribution</i>	<i>Created product is interesting, new, and/or helpful, making an original contribution that includes identifying a previously unknown problem, issue, or purpose</i>	<i>Created product is interesting, new, and/or helpful, making an original contribution for its intended purpose (e.g., solving a problem or addressing an issue)</i>	<i>Created serves its intended purpose (e.g., solving a problem or addressing an issue)</i>	<i>Created does not serve its intended purpose (e.g., solving a problem or addressing an issue)</i>

Brookhart (2013)

3. Deskriptor

E. Langkah-langkah Merancang Rubrik

Untuk merancang sebuah rubrik, guru dapat memulainya dengan membaca literatur tentang rubrik untuk memahami prinsip-prinsip pembuatan rubrik yang baik. Selain itu, guru dapat melihat sumber-sumber *online* untuk menemukan rubrik yang telah ada atau dibuat orang lain. Tidak perlu untuk membuat sendiri rubrik; guru bisa mengadaptasi rubrik yang ada. Guru cukup merevisi sesuai dengan kebutuhan. Setelah guru menemukan model rubrik, guru dapat menyesuaikan kriteria, skala penilaian, dan indikator dengan kebutuhan penilaian kelas.

Untuk merancang sendiri rubrik atau mengadaptasi rubrik yang telah ada, proses dimulai dengan mengidentifikasi komponen rubrik dasar: kriteria kinerja, skala peringkat, dan indikator dari kinerja. Berikut langkah-langkah yang dapat dilakukan.

1. Tentukan Kriteria

Cari tahu kriteria apa yang benar-benar penting dari kualitas tugas yang sedang dikerjakan siswa. Apakah itu esai, proyek, cerita digital atau esai, atau presentasi, apa yang guru inginkan dari proses atau produk akhir?

- a. Sebutkan semua kriteria yang mungkin guru ingin siswa tunjukkan dalam tugas. Masukkan kriteria untuk proses menciptakan produk dan kualitas produk.
- b. Tentukan kriteria mana yang “tidak bisa dinegosiasikan.” Idealnya, rubrik tersebut akan memiliki tiga hingga lima kriteria kinerja. Jika kesulitan menentukan, prioritaskan kriteria dengan bertanya: Apa hasil pembelajaran pokok bahasan ini? Hasil pembelajaran mana yang akan dicantumkan dalam rubrik? Keterampilan mana yang penting pada tingkat kompeten atau kemahiran yang diharapkan? Seberapa pentingkah penyelesaian keseluruhan tugas atau proyek?

2. Kembangkan Skala Peringkat Atau Rentang Level

Skala penilaian dapat mencakup label numerik atau deskriptif. Biasanya, skala peringkat terdiri dari jumlah level kinerja yang genap. Tampilkan skala peringkat Anda mulai dari kiri dengan yang tertinggi. Buat indikator yang ada di semua level kinerja.

3. Kembangkan Indikator Kualitas

- a. Buat indikator yang ada di semua level kinerja.
- b. Pastikan ada kesinambungan dalam perbedaan antara setiap kriteria satu dengan lainnya untuk melebihi vs bertemu, dan bertemu vs tidak memenuhi harapan.
- c. Lakukan pengeditan terhadap indikator untuk memastikan bahwa level mencerminkan varians dalam kualitas dan bukan pentingnya kriteria. Pastikan bahwa indikator mencerminkan interval yang sama di sepanjang skala.

F. Me-review Rubrik

Saat guru mengadopsi, mengadaptasi atau membuat sendiri rubrik, guru harus meninjau rubrik yang ada atau yang telah dibuat untuk menentukan apakah rubrik tersebut efektif dan sesuai dengan kebutuhan. Untuk memandu guru, Universitas Minnesota membuat panduan pertanyaan-pertanyaan seperti berikut ini.

1. Apakah karakteristik setiap tingkat kinerja jelas?
2. Apakah siswa dapat menilai sendiri dengan memiliki deskriptor?
3. Akankah deskriptor memberi siswa informasi yang cukup untuk mengetahui apa yang mereka perlu tingkatkan?
4. Apakah rubrik secara memadai mencerminkan kisaran level di mana siswa dapat benar-benar melakukan tugas yang diberikan?
5. Apakah kriteria pada setiap level didefinisikan dengan jelas untuk memastikan bahwa penilaian itu akurat, tidak bias, dan konsisten? Bisakah beberapa instruktur menggunakan rubrik dan menilai kinerja siswa dalam rentang yang sama?
6. Apakah rubrik mencerminkan proses dan produk?
7. Apakah semua kriteria sama pentingnya, atau satu variabel lebih kuat dari yang lain?
8. Apakah bahasa yang digunakan deskriptif untuk siswa untuk menentukan apa yang diukur dalam metode kualitatif dan kuantitatif.

Hal-hal lain yang juga penting untuk dilakukan dalam merancang rubrik. Rubrik perlu diujicobakan, atau diuji di lapangan, untuk memastikan mereka mengukur variabel yang dimaksudkan oleh

perancang. Rubrik dapat didiskusikan dengan siswa untuk menciptakan pemahaman tentang apa yang diharapkan dari siswa. Rubrik mencantumkan ekspektasi kinerja siswa yang selaras dengan pelajaran konseptual atau unit dikirimkan. Siswa seharusnya tidak diharapkan untuk melakukan apa yang sebelumnya tidak diajarkan atau ditunjukkan kepada mereka.

G. Menggunakan Rubrik

Bagaimana menggunakan rubrik agar memberikan manfaat yang maksimal terhadap pembelajaran? Berikut ini adalah contoh yang bisa guru lakukan. Salah satunya adalah dengan melibatkan siswa secara aktif.

1. Perhatikan model rubrik: Perhatikan kepada siswa contoh-contoh pekerjaan yang baik dan yang tidak terlalu baik. Identifikasi karakteristik yang membuat yang baik menjadi baik dan yang buruk menjadi buruk.
2. Buatlah daftar kriteria: Gunakan model yang sudah ditetapkan untuk membuat daftar apa saja yang penting untuk sebuah karya yang berkualitas.
3. Mengartikulasikan gradasi kualitas: Jelaskan level kualitas terbaik dan terburuk, kemudian isi level menengah.
4. Berlatihlah menggunakan model: Mintalah siswa menggunakan rubrik untuk mengevaluasi model yang guru berikan pada Langkah 1.
5. Minta siswa melakukan penilaian sendiri dan teman sebaya: Berikan tugas kepada siswa. Saat mereka bekerja, hentikan mereka sesekali untuk penilaian sendiri dan rekan.
6. Merevisi: Selalu beri siswa waktu untuk merevisi pekerjaan mereka berdasarkan umpan balik yang mereka dapatkan di Langkah 5.
7. Gunakan penilaian guru: Gunakan rubrik yang sama yang digunakan siswa untuk menilai pekerjaan mereka sendiri. Langkah 1 mungkin diperlukan hanya ketika Anda meminta siswa untuk terlibat dalam tugas yang mereka tidak kenal. Langkah 3 dan 4 berguna tetapi memakan waktu; Anda dapat melakukannya sendiri, terutama ketika Anda telah menggunakan rubrik untuk sementara waktu. Kelas yang berpengalaman dalam penilaian berbasis rubrik dapat

merampingkan proses sehingga dimulai dengan kriteria daftar, setelah itu guru menuliskan gradasi kualitas, memeriksa mereka dengan siswa, membuat revisi, kemudian menggunakan rubrik untuk diri sendiri, teman sebaya, dan penilaian guru.

Latihan

1. Tinjaulah kembali keunggulan dan kelemahan empat jenis rubrik yang telah dibahas dalam bab ini! Buatlah tabel yang memperlihatkan bagaimana jenis rubrik dapat digunakan dalam menilai berbagai tugas atau format tes, maupun tujuan pembelajaran, di kelas Anda!
2. Tinjaulah rubrik yang pernah Anda buat atau gunakan dan tentukan apakah rubrik tersebut sudah memenuhi prinsip-prinsip rubrik yang baik!

[Halaman ini sengaja dikosongkan]

Bab 9

Menilai Keterampilan Berpikir Tingkat Tinggi (HOTS)

Objek Pembelajaran

- A. Prinsip Umum dalam Menilai *Higher Order Thinking Skills* (HOTS)
- B. Strategi dalam Memberikan Umpan Balik dan Penskoran Tugas dalam HOTS

Menggunakan kemampuan berpikir kritis untuk memecahkan sebuah masalah sangat dianjurkan bagi siswa dalam proses pembelajaran di era industri 4.0 ini. *High Order Thinking Skills* (HOTS) atau dalam Bahasa Indonesia disebut kemampuan berpikir tingkat tinggi, merupakan suatu kemampuan berpikir yang sedang digalakkan oleh pemerintah untuk meningkatkan kualitas pendidikan. Brookhart (2010) mengklasifikasikan pengertian HOTS ke dalam tiga kategori: (1) pengertian *higher order thinking skills* dalam hal transfer, (2) pengertiannya dalam hal berpikir kritis, (3) pengertian *higher order thinking skills* dalam hal pemecahan masalah.

Dalam kategori pertama (*transfer*) dengan menggunakan HOTS siswa tidak hanya mampu mengingat pembelajaran yang telah diajarkan tetapi mereka juga harus bisa memahami serta menggunakan pembelajaran tersebut. Maksudnya adalah siswa dapat mengaplikasikan pengetahuan dan kemampuannya selagi mereka mempelajari hal yang baru. Hal baru maksudnya sesuatu yang belum siswa ketahui sebelumnya. Caranya yaitu menghubungkan materi yang sedang siswa pelajari dengan apa yang mereka bayangkan agar tercipta sebuah asosiasi. Contohnya yaitu ketika siswa taman kanak-kanak mempelajari tentang angka, mereka terkadang mengasosiasikan angka-angka tersebut dengan bentuk benda yang mereka temui sehari-hari. Misal, angka dua (2) bentuknya seperti bebek, angka 7 seperti ular, angka 1 seperti tongkat, dan lain-lain. Kemampuan seperti itu yang disebut dengan *higher-order thinking skills*.

Seperti yang kita tahu bahwa mengajar adalah proses pentransferan ilmu kepada siswa di mana hal tersebut merupakan sebuah tujuan umum dari pendidikan. Transfer ilmu tersebut ditujukan agar siswa dapat menerapkannya di kehidupan nyata tanpa adanya lagi bimbingan dan arahan dari guru. Tentu saja mereka harus menggunakan kemampuan berpikirnya sendiri untuk menjalani semua setelah lepas dari bimbingan dan arahan gurunya. Oleh karena itu, *higher order thinking skills* sangat efisien diterapkan dalam mendidik dan membimbing siswa dalam belajar.

Selanjutnya, pada kategori yang kedua adalah berpikir kritis (*critical thinking*), HOTS dilakukan agar siswa mampu berpikir yang mana mereka harus fokus dalam memutuskan apa yang mereka pikirkan dan lakukan. Maksudnya adalah siswa mampu membuat pendapat yang bijak atau mereka dapat menciptakan sebuah pendapat yang kritis dengan landasan yang kuat. Jadi, pada kategori ini tujuan digunakannya HOTS dalam sebuah pengajaran yaitu agar siswa mampu berpendapat, merefleksikan pendapatnya, dan membuat sebuah keputusan yang kritis.

Kategori terakhir yaitu pemecahan masalah (*problem solving*). Tentunya dalam belajar untuk mencapai tujuan pembelajarannya siswa acap kali menemukan banyak sekali hambatan di mana siswa tidak secara otomatis menemukan solusi untuk keluar dari kesulitannya tersebut. Pada situasi seperti itu, siswa harus menggunakan salah satu cara dari HOTS yaitu *problem solving*. *Problem solving* merupakan proses memecahkan sebuah masalah yang dihadapi siswa dengan cara

mengingat materi atau informasi yang telah diberikan guru, belajar dengan pemahaman sendiri, mengevaluasinya dengan kritis, membuat cara-cara yang kreatif, dan mengomunikasikan seefektif mungkin. Dengan mempertimbangkan ketiga kategori tersebut, siswa dapat belajar dengan baik dan meningkatkan kemampuan berpikirnya dalam belajar tentunya.

Dengan adanya tiga kategori atau pendekatan terhadap HOTS tersebut, banyak sekolah-sekolah yang mencoba mengaplikasikannya dalam proses belajar mengajar. Tentu dalam proses belajar tidak pernah luput dengan istilah penilaian. Dalam bab ini kita akan membahas bagaimana menilai atau melakukan asesmen untuk metode pembelajaran HOTS ini. Apa saja yang akan dinilai? Dan bagaimana cara menilainya?

A. Prinsip Umum dalam Menilai *Higher Order Thinking (HOTS)*

Apakah Anda masih ingat prinsip dasar sebuah penilaian? Baiklah, seperti yang kita telah bahas sebelumnya, bahwa kita harus menentukan apa yang ingin kita nilai, membuat tugas atau tes untuk melihat sejauh mana kemampuan dan pengetahuan siswa, dan membuat laporan sebagai bukti untuk menjelaskan sudah sejauh mana kemampuan dan pemahaman siswanya. Ya tentu itu merupakan prinsip dasar dari sebuah penilaian. Pada dasarnya prinsip dasar dalam menilai HOTS juga bersumber pada prinsip dasar penilaian, namun di bawah ini ada tiga tambahan lain, di antaranya seperti yang disampaikan oleh Brookhart (2010).

1. Menyajikan sesuatu kepada siswa untuk dipikirkan, biasanya berbentuk teks, materi, visual, skenario, atau semacam masalah atau kasus yang berhubungan dengan konteks pembelajaran.
2. Memberikan materi-materi atau bahan ajar yang asing diketahui oleh siswa tentu yang masih berhubungan dengan pelajaran yang sedang diajar.
3. Membedakan antara level kesulitan (mudah vs sulit) dan level berpikir (*Lower-order thinking vs Higher-order thinking*).

Tentu ketiga prinsip tersebut sangat penting agar proses pembelajaran menggunakan HOTS berjalan dengan baik. Tanpa adanya prinsip penilaian HOTS mungkin dalam penerapan di kelas akan gagal.

Untuk mengetahui lebih detail mengenai apa yang harus guru dan siswa lakukan dalam menerapkan prinsip di atas, mari kita bahas lebih lanjut mengenai ketiga prinsip penilaian dari HOTS.

Prinsip pertama yaitu menyajikan materi atau bahan kepada siswa untuk dipikirkan. Dengan kata lain, guru disarankan untuk menggunakan materi pengantar. Menggunakan materi pengantar maksudnya adalah memberikan materi kepada siswa untuk ditelaah. Misalnya guru memberikan sebuah tes membaca berisi teks tentang “Energi Nuklir Akan Mengusung Energi Listrik”. Dalam penilaian HOTS siswa tidak diizinkan untuk menghafal bacaan tersebut, namun yang dianjurkan adalah memahami bacaan tersebut. Seperti yang kita tahu jika kita menghafalkan sesuatu kemungkinan besar untuk lupa terhadap materi pada saat tes lebih besar. Berbeda jika kita memahaminya dengan sungguh-sungguh, bentuk soal seperti apa pun pasti kita akan bisa menjawabnya.

Prinsip yang kedua yaitu memberikan bahan ajar yang baru, maksudnya bahan yang asing bagi siswa namun masih dalam konteks materi yang diajarkan. Memang terlihat kontras dengan sifat dasar dari sebuah penilaian di mana seharusnya guru menggunakan materi yang familiar untuk mengukur kemampuan siswa. Perlu kita semua ketahui bahwa tujuan dari prinsip ini tentu supaya siswa berpikir kritis bukan menghafal materinya. Itulah mengapa guru dianjurkan untuk membiasakan siswanya dengan materi-materi belajar yang baru yang memang siswa belum pernah mencobanya agar mereka terbiasa berpikir kritis dalam menghadapi sebuah masalah. Contohnya ketika akan ada penilaian akhir atau penilaian sumatif, siswa bisa menggunakan kemampuan HOTS ini untuk mengerjakan soal yang memang belum pernah mereka hadapi sebelumnya.

Kita masuk prinsip ketiga yaitu mengatur tingkat kerumitan dan tingkat kesulitan kognitif secara terpisah. Perlu diketahui bahwa dalam menilai kedua tingkat tersebut harus menggunakan pertanyaan atau tugas dengan sistem HOTS. Seperti yang telah dibahas sebelumnya bahwa dalam tugas atau soal HOTS ditujukan untuk siswa agar dapat berpikir kritis untuk meningkatkan kemampuan berpikirnya. Namun banyak kesalahpahaman yang timbul bahwa menggunakan *recall* itu mudah dan menggunakan HOTS itu sulit dan menggiring siswa untuk mendapatkan nilai yang jelek. Maka dari itu, perlu kita sebagai guru

untuk mengatur atau menyesuaikan tingkat kerumitan dan kesulitan soal pada saat menerapkan HOTS yang akan diujikan kepada siswa. Di bawah ini merupakan contoh tingkat kerumitan dan kesulitan soal pada saat menggunakan *recall* dan HOTS.

Tabel 9.1 Perbedaan Tingkat Kerumitan dan Kesulitan Soal Antara *Recall* dan HOTS

	<i>Easy</i>	<i>Difficult</i>
Recall	<i>Who is the main character in The cat in the Hat?</i>	<i>Name all the characters in Hamlet</i>
Higher-Order Thinking	<i>Why do you think the Cat Cleaned up the house on his way out, before Mother got home?</i>	<i>Hamlet wrestles with a major question in his soliloquy, “O, that this too, too solid flesh would melt” in Act 1, Scene 2, Lines 131-161. What is the question in his mind, and how do you think he resolves it by the end of his soliloquy? State your interpretation of his major question and his resolution and use evidence from the speech to support it.</i>

Sumber: Diambil dari Buku *How to Assess High-Order Thinking Skill* (Brookhart, 2010)

B. Strategi dalam Memberikan *Feedback* dan Penskoran Tugas dalam HOTS

Menurut Brookhart (2010), ada dua cara untuk menginterpretasikan jawaban siswa ketika mengerjakan soal atau tugas, yaitu memberikan komentar pada pekerjaannya dan memberikan skor. Kemudian ada hal penting yang harus diperhatikan oleh guru ketika membuat tes menggunakan HOTS, yaitu menampilkan pada lembar soal mengenai kriteria (yang terlihat pada contoh soal HOTS di atas). Ternyata kriteria tersebut menjadi fondasi untuk memberikan *feedback* kepada siswa atau sebagai landasan rubrik penilaian. Namun semuanya tergantung atau sesuai kebutuhan bagaimana kamu ingin meng-*assess* siswa. Terpenting kriteria-kriteria tersebut dapat atau sesuai target dari pembelajarannya.

1. Penilaian *Formative Higher-Order Thinking Skills*

Salah satu cara yang ampuh untuk menilai atau meng-*assess high-order thinking skills*, yaitu dengan melakukan observasi atau diskusi secara langsung mengenai penalaran siswa (Brookhart, 2010: 31). Bagaimana tidak, dengan melakukan obrolan-obrolan dengan siswa kita bisa mendapatkan dan melihat *progress* kemampuan siswa (formatif). Tentu obrolan yang diadakan masih berhubungan dengan materi pembelajaran. Selain melakukan observasi dan diskusi, seperti biasanya, guru juga bisa menggunakan *feedback* tertulis.

Misalkan ketika guru ingin mengetahui sejauh mana kemampuan *vocabulary* siswa, guru mengadakan forum diskusi di kelas untuk mengidentifikasi mengapa susah untuk memahami kosakata dalam Bahasa Inggris, menghafalkannya, dan menerapkannya pada praktik penggunaan bahasa. Contoh lain ketika guru memberikan *feedback* tertulis pada saat siswa mengerjakan tugas *writing*: “Ide dari tulisan kamu bagus dan menarik, namun yang perlu diketahui dan harus ditingkatkan ialah bagaimana kamu mengelaborasi setiap paragraf yang ada dan kamu juga harus memberikan ilustrasi dari setiap gagasan yang telah kamu tulis”.

Komentar pada pekerjaannya dan memberikan skor. Kemudian ada hal penting yang harus diperhatikan oleh guru ketika membuat tes menggunakan HOTS, yaitu menampilkan pada lembar soal mengenai kriteria (yang terlihat pada contoh soal HOTS di atas). Ternyata kriteria tersebut menjadi fondasi untuk memberikan *feedback* kepada siswa atau sebagai landasan rubrik penilaian. Namun semuanya tergantung atau sesuai kebutuhan bagaimana kamu ingin meng-*assess* siswa. Terpenting kriteria-kriteria tersebut dapat atau sesuai target dari pembelajarannya.

2. Penilaian *Summative Higher-Order Thinking*

Pada penilaian sumatif, sebuah soal atau tugas yang dituntut untuk mengukur HOTS dapat dirusak oleh skema penilaian yang hanya memberi skor pada fakta yang dilaporkan saja (Brookhart, 2010). Maksudnya, kemampuan *higher-order thinking skills* siswa tidak akan tampak jika guru tidak menilai kualitas penalaran siswa. Untuk membuat skema penilaian yang bisa mengukur HOTS siswa, guru harus membuat kriteria dari rubrik yang dikembangkan. Seperti yang kita tahu rubrik pada umumnya hanya menghitung nomor yang benar dari jawaban siswa. Namun untuk

menentukan skor HOTS siswa, guru harus mengganti skema penilaian yang ada menjadi sebuah latihan untuk mendapatkan nilai tersebut. Contohnya dalam soal pilihan ganda, di mana setiap nomor yang benar memiliki skor 1 dan yang salah 0. Sedangkan untuk menghasilkan skor yang menandakan siswa menggunakan *high-order thinking skill*, soalnya harus didesain sedemikian rupa agar kemampuan *high-order thinking skill*-nya benar-benar harus dijawab.

Dalam rubrik *constructed-response* dan soal esai, kita harus mengawali dengan menentukan kriteria atau jenis penalaran apa yang akan dinilai. Misalnya “Apakah siswa menimbang fakta-fakta sebelum membuat keputusan?” atau “Apakah siswa mengukur tingkat kredibilitas sumbernya sebelum menentukan jawabannya?”.

Skema penilaian yang terakhir yang harus diperhatikan dalam menilai kemampuan *higher-order thinking* siswa, yaitu skema penilaian pada penilaian performansi. Analisis rubrik sering digunakan untuk menilai *performance assessment*, *papers*, dan proyek. Dalam hal ini, guru dapat membuat sendiri rubrik penilaiannya atau mencari referensi dari internet dan kurikulum yang dipakai. Dalam memilih referensi rubrik di internet kita harus berhati-hati dan selalu memastikan bahwa rubrik tersebut sesuai dengan apa yang kita nilai. Selain itu, kita juga dianjurkan untuk membuat atau menulis rubrik yang mampu mendeskripsikan kualitas berpikir dibandingkan yang hanya menghitung angka. Di bawah ini merupakan contoh rubrik yang cocok digunakan untuk membuat penilaian pada penilaian *performance (problem solving)*.

	4	3	2	1
Identifying Problems	I anticipate different kinds of problems in complicated projects and think of ways to solve them before they happen.	I anticipate problems in complicated projects and think of ways to avoid them.	Sometimes I anticipate problems in complicated projects, but I usually do not think of ways to avoid them.	I have a hard time anticipating problems before they happen.
Identifying Relevant Information	I clearly identify important information needed to solve complex problems.	I separate important information from unimportant information when solving problems.	I sometimes need help identifying important information when solving problems.	I often miss important information when solving problems.
Analyzing Problems	I carefully analyze all the characteristics of a problem before I begin to solve it.	I think carefully about a problem before I try to solve it.	Sometimes, I start to solve a problem before I think about it.	I usually start to solve a problem before I think about it.
Using Strategies	I use the strategies and tools I learned and subject-area knowledge to solve problems, even when they are not school assignments.	I use subject-area strategies, tools, and knowledge to solve problems.	If I am reminded, I use subject-area strategies, tools, or knowledge to solve problems.	I do not use subject-area strategies, tools, or knowledge to solve problems.
Reflecting	I reflect on my problem solving processes, evaluate how well they are working, and make changes when necessary.	I reflect on my problem solving processes by thinking about what I did well and what I can do better.	I sometimes go back and think about how I solved a problem after I am finished.	I hardly ever go back and think about how I solved a problem after I am finished.

Sumber: Diambil dari www.pinterest.com (*Educational Technology and Mobile Learning: Awesome Problem-Solving Rubric for Teachers*).

Dapat kita lihat di atas merupakan rubrik yang digunakan untuk menilai seberapa jauh siswa dapat melakukan *problem solving*. Rubrik di atas bisa masih bersifat umum dan belum fokus terhadap sebuah pembelajaran contohnya Bahasa Inggris, Matematika, dan Sains. Tentu rubrik tersebut dapat digunakan dan dimodifikasi sesuai pelajaran yang akan kita uji, contohnya untuk menilai *skill* dalam Bahasa Inggris.

Pada rubrik di atas, terlihat mempunyai lima kategori, yang pertama adalah mengidentifikasi masalah. Kategori ini berisi tentang bagaimana siswa mengantisipasi sebuah masalah pada tugas atau proyek yang diberikan guru. Ada empat beban skor yang mendeskripsikan bagaimana siswa dalam mengantisipasi sebuah masalah. Kemudian kategori kedua adalah mengidentifikasi informasi yang relevan. Pada kategori ini, siswa dinilai bagaimana mereka menggunakan informasi untuk memecahkan sebuah masalah. Kategori ketiga yaitu menganalisis masalah. Siswa dinilai bagaimana mereka menganalisis masalah atau tugas yang diberikan oleh guru. Selanjutnya, penggunaan strategi dalam memecahkan sebuah masalah. Pada kategori ini siswa dilihat apa dan bagaimana mereka menggunakan strateginya untuk memecahkan sebuah masalah. Kategori yang terakhir yaitu refleksi atau bisa disebut kroscek. Di dalam kategori guru menilai siswa bagaimana mereka mengkroscek apa yang telah dilakukan dalam memecahkan sebuah masalah.

Itulah tadi cara dasar melakukan penilaian untuk mengukur HOTS. *Higher Order Thinking* merupakan metode yang mampu menstimulasi siswa untuk berpikir kritis dan kreatif, jadi sangat penting bagi kita untuk tahu cara menggunakan dan menilai metode ini. Perlu pembiasaan yang intens agar HOTS dapat digunakan secara optimal. Pemilihan materi yang *out of the box* namun menarik untuk dipelajari merupakan kunci untuk memperkenalkan HOTS kepada para siswa.

Latihan

1. Berikan contoh bentuk penilaian HOTS dalam asesmen formatif!
2. Ada yang mengatakan bahwa soal-soal yang mengandung HOTS selalu sulit/rumit. Apakah Anda setuju dengan pernyataan di atas? Kemukakan alasannya!

[Halaman ini sengaja dikosongkan]

Bab 10

Umpan Balik yang Efektif

Objek Pembelajaran

- A. Konsep Umpan Balik (*Feedback*)
- B. Strategi dan Konten Umpan Balik
- C. Teknik Mengajarkan Siswa Memanfaatkan Umpan Balik

A. Konsep Umpan Balik (*Feedback*)

Umpan balik adalah komponen penting dari proses penilaian formatif. Penilaian formatif memberikan informasi kepada guru dan siswa tentang bagaimana pencapaian siswa terhadap tujuan pembelajaran di kelas. Memberikan umpan balik yang baik adalah salah satu keterampilan yang harus dikuasai guru sebagai bagian dari penilaian formatif yang baik.

Dampak umpan balik akan sangat signifikan jika umpan balik dilakukan dengan baik. Dua keunggulan, yaitu menyorot aspek kognitif dan sekaligus aspek motivasi pada saat yang sama. Faktor kognitif berarti umpan balik yang baik memberi siswa informasi yang mereka butuhkan sehingga mereka dapat memahami di mana mereka berada

dalam pembelajaran mereka dan apa yang harus dilakukan selanjutnya. Saat siswa merasa memahami apa yang harus dilakukan dan alasannya, siswa akan termotivasi untuk meningkatkan pencapaian mereka dan menjadi pembelajar yang mandiri.

Umpan balik yang baik bisa menciptakan lingkungan dan budaya penilaian kelas yang baik pula. Umpan balik yang baik membuat siswa menerima kritik membangun dari guru sebagai sesuatu yang positif dan memahami bahwa latihan terus-menerus adalah proses belajar. Sebaliknya, jika budaya yang tumbuh adalah budaya untuk selalu memperbaiki kesalahan, kritik dari guru akan dianggap negatif.

B. Strategi dan Konten Umpan Balik

Strategi yang digunakan guru dalam memberikan umpan balik sangat berpengaruh terhadap keefektifan umpan balik. Demikian pula dengan konten umpan balik itu sendiri.

1. Strategi Umpan Balik

Menurut Brookhart (2008), ada empat hal yang perlu diperhatikan dalam memberikan umpan balik. Umpan balik yang diberikan akan bervariasi tergantung dari waktu, jumlah, moda, dan audiensnya. Waktu berkaitan dengan kapan dan bagaimana umpan balik diberikan. Beberapa saran yang bisa dilakukan guru adalah berikanlah umpan balik untuk sesuatu yang berkaitan dengan pengetahuan atau fakta. Guru bisa menunda umpan balik untuk tugas yang berkaitan dengan penjelasan yang komprehensif tentang pemikiran siswa. Jangan menunda memberikan umpan balik agar tidak kehilangan momentum. Berikan umpan balik sesering yang guru bisa terutama untuk tugas-tugas berupa proyek besar.

Aspek waktu yang kedua berkaitan dengan jumlah atau porsi umpan balik. Beberapa hal yang diberi umpan balik? Seberapa detail umpan balik untuk setiap hal? Agar efektif, lakukanlah prioritas, pilihlah dan berilah umpan balik pada poin atau hal yang paling penting saja. Pilihlah hal yang berkaitan dengan tujuan utama pembelajaran dan pertimbangkan tingkat perkembangan masing-masing siswa.

Moda umpan balik juga bervariasi. Bisa verbal, tertulis, maupun berupa peragaan atau visual. Guru harus memutuskan apakah umpan

balik yang berupa komentar sambil lalu sudah cukup memadai atautkah diskusi khusus yang diperlukan. Umpan balik yang interaktif oleh guru dengan siswa secara individu adalah praktik yang terbaik jika memungkinkan. Jika tugas yang dikerjakan siswa berbentuk tugas tertulis, berilah umpan balik tertulis pula. Umpan balik berupa peragaan dapat dilakukan jika tugas yang dikerjakan siswa berkaitan dengan bagaimana cara melakukan sesuatu dan siswa masih mengalami kesulitan dalam hal ini.

Aspek audiens juga memengaruhi strategi guru dalam memberikan umpan balik. Apakah umpan balik diberikan secara individual atau secara berkelompok. Umpan balik secara individu membuat siswa merasa dihargai. Umpan balik secara berkelompok patut dilakukan jika mayoritas siswa belum memahami tugas yang mereka kerjakan.

Dari paparan di atas dapat disimpulkan bahwa strategi pemberian umpan balik tergantung dengan situasi dan kondisi. Tidak ada rumus khusus untuk melakukan umpan balik terbaik untuk setiap siswa, semua tergantung dari konteksnya.

2. Konten Umpan Balik

Selain strategi, tentu saja guru harus memikirkan konten dari umpan baliknya. Seperti halnya strategi, konten dari umpan balik juga bervariasi tergantung dengan beberapa aspek. Brookhart menyebut lima aspek seperti fokus, perbandingan, fungsi, valensi, kejelasan, spesifikasi, dan nada dari isi umpan balik.

Fokus dari sebuah umpan balik yang dilakukan guru dapat berfokus kepada tugas akhir siswa, pada proses yang digunakan siswa dalam mengerjakan tugas tersebut, pada kemandirian siswa dalam mengerjakan tugas atau kepada diri siswa itu sendiri. Umpan balik yang baik dalam hal ini sebaiknya dilakukan dengan cara berikut. Jika memungkinkan, fokuskan umpan balik untuk menjelaskan tugas dan prosesnya serta hubungan keduanya. Berilah komentar tentang kemandirian siswa dalam mengerjakan tugas sehingga dapat menumbuhkan rasa percaya diri siswa. Dan yang tak kalah penting adalah hindari memberi komentar yang personal.

Umpan balik juga bisa difokuskan pada perbandingan yaitu membandingkan hasil kerja siswa terhadap kriteria kerja atau tugas yang

baik, membandingkan dengan hasil siswa lain, dan membandingkan hasil siswa yang bersangkutan dengan hasil yang dicapai sebelumnya.

Dalam hal ini, guru disarankan menggunakan umpan balik yang merujuk kepada kriteria untuk memberikan informasi yang berfokus pada tugas/PR siswa. Jika guru ingin memberikan umpan balik dengan membandingkan satu siswa dengan siswa yang lain, dan memberikan informasi tentang proses dan usaha yang siswa lakukan gunakan perbandingan norma. Bentuk lain dari umpan balik dalam kategori ini adalah siswa melakukan umpan balik terhadap diri sendiri bagi siswa yang belum berhasil agar mereka fokus untuk melihat seberapa banyak kemajuan yang mereka capai, dan bukan pada seberapa jauh mereka tertinggal dalam mencapai tujuan pembelajaran.

Konten sebuah umpan balik juga dilihat dari fungsinya. Jika isi umpan balik tersebut adalah penjelasan, maka konten umpan balik guru harus berfungsi sebagai penjelasan. Jika umpan balik dimaksudkan sebagai bentuk evaluasi, hindari untuk menghakimi hasil kerja siswa.

Aspek lain dari isi umpan balik adalah apakah isinya positif atau negatif. Untuk menggambarkan bahwa siswa telah melakukan tugas dengan baik, gunakan komentar positif. Jika guru memberikan umpan balik negatif, harus diimbangi dengan komentar positif tentang apa yang bisa ditingkatkan.

Kejelasan sebuah umpan balik juga masuk dalam kategori isi. Umpan balik harus jelas dan dapat dipahami oleh siswa. Untuk itu, guru harus menggunakan kosakata dan konsep yang mudah dimengerti. Guru juga harus menyesuaikan jumlah isi umpan balik dengan tingkat perkembangan masing-masing siswa.

Hal lain yang terkait dengan kejelasan ini adalah seberapa spesifik umpan balik yang diberikan oleh guru; sangat spesifik, spesifik, atau umum saja. Dalam hal ini, disarankan agar guru menyesuaikan tingkat kespesifikan umpan balik pada siswa dan pada tugas yang dikerjakan siswa. Buatlah umpan balik cukup spesifik agar siswa tahu apa yang harus dilakukan tapi tidak terlalu spesifik yang membuat siswa tidak perlu melakukan apa pun. Dalam memberikan umpan balik, guru sebaiknya mengidentifikasi kesalahan atau jenis kesalahan, namun hindari mengoreksi semua kesalahan tersebut. Biarkan siswa berlatih untuk mengoreksi kesalahan mereka.

Dalam memberikan umpan balik, guru juga harus mempertimbangkan implikasi dari umpan balik. Nada dari umpan balik akan memengaruhi apa yang siswa dengar. Oleh karena itu, guru harus memilih kata-kata yang santun namun tepat sasaran dan kata-kata yang menempatkan siswa sebagai subjek. Tak kalah pentingnya, gunakanlah kata-kata yang merangsang siswa untuk berpikir.

C. Teknik Mengajarkan Siswa Memanfaatkan Umpan Balik

Pertanyaan yang biasa diajukan setelah guru menjalankan tugasnya memberikan umpan balik adalah “Apa yang dilakukan siswa dengan umpan balik yang mereka terima?” Beberapa penelitian mengungkapkan bahwa banyak siswa yang tidak melakukan apa pun setelah menerima umpan balik karena mereka tidak tahu apa yang harus dilakukan. Oleh karena itu, guru perlu mengajarkan siswa bagaimana merespons umpan balik sebagai alat pembelajaran. Secara umum, guru bisa menggunakan berbagai strategi untuk mengajarkan siswa memanfaatkan umpan balik termasuk di antaranya memberikan contoh langsung, mengajarkan cara menilai diri sendiri dan teman sebaya.

Teknik yang pertama adalah guru memberikan contoh. Teknik ini merupakan salah satu teknik terbaik untuk mengajarkan siswa. Guru dapat memberi contoh cara memberi dan menggunakan umpan balik sebagai bagian dari pelajaran. Guru juga dapat memberi contoh untuk terbuka terhadap kritik membangun yang diberikan oleh guru atau teman sebaya. Siswa harus mengharapkan kritik dan menganggapnya sebagai kesempatan untuk belajar.

Teknik yang kedua adalah mengajarkan siswa untuk melakukan penilaian diri sendiri dan teman sebaya. Pendekatan ini mengajarkan mereka dari mana umpan balik berasal. Mereka akan mempelajari strategi pada saat yang sama ketika mereka belajar bagaimana meningkatkan *performance* mereka, tugas menulis, pemecahan masalah atau apa pun yang mereka kerjakan. *Self assessment* meningkatkan minat siswa dalam umpan balik karena mereka melakukan sendiri; menjawab pertanyaan mereka sendiri dan membantu mereka mengembangkan keterampilan menjadi pembelajar yang mandiri.

Salah satu teknik lain yang dapat membantu siswa adalah membuat rubrik yang sederhana dan mudah dimengerti oleh siswa. Untuk

membantu siswa memahami kriteria dan memahami umpan baliknya, guru bisa melakukan beberapa hal. Misalnya, guru menyiapkan latihan sebelum memberikan tugas kepada siswa. Guru menjelaskan tentang tugas yang akan dikerjakan oleh siswa. Kemudian membagi siswa ke dalam beberapa kelompok dan membagikan rubrik yang dibuat oleh guru. Siswa diminta untuk menerjemahkan rubrik tersebut ke dalam bahasa mereka sendiri. Sambil berkeliling kelas, guru memberi komentar untuk membantu siswa dalam diskusi mereka. Kemudian rubrik hasil terjemahan siswa dikumpulkan. Manfaat dari latihan seperti ini adalah siswa memahami kriteria rubrik, dan karena itu mereka diharapkan lebih mandiri dalam mengerjakan tugas.

Contoh latihan lain yang bisa dilakukan adalah melakukan *review* terhadap hasil tes yang sudah dinilai untuk melihat apa yang mereka ketahui dan tidak ketahui, dan merencanakan strategi untuk meningkatkan pengetahuan mereka pada mata pelajaran tersebut. Ajaklah siswa berdiskusi tentang berbagai alasan mengapa hasil tes siswa dianggap salah oleh penilai. Menurut Brookhart (2008) dan Andrade (2013) ada beberapa jenis kesalahan yang dibuat siswa.

1. Siswa tahu jawabannya tetapi karena alasan tertentu tidak menandainya dengan benar.
2. Siswa seharusnya tahu jawabannya tetapi membaca pertanyaan itu dengan sangat cepat, kata penting dilewati atau kesalahan serupa yang berhubungan dengan kebiasaan.
3. Siswa mengira dia tahu jawabannya tetapi salah dalam pemahamannya.
4. Kurang pengetahuan.

Untuk memahami tipe kesalahan yang dibuat tidaklah mudah. Brookhart (2008) memberikan contoh tabel berikut untuk membantu siswa.

Pertanyaan	Jawaban Benar	Apakah jawaban saya salah? Jika ya, mengapa?	Apa yang harus saya lakukan?
1			
2			
3			
4			

5			
6			
dst			

Lembar umpan balik tes (diadaptasi dari Brookhart, 2008)

Teknik lainnya untuk mengajarkan siswa adalah dengan menetapkan target dan kriteria pembelajaran yang jelas. Penugasan dan rubrik yang berkualitas dan jelas memungkinkan umpan balik yang baik. Jika siswa memperhatikan fokus menghasilkan karya yang berkualitas proses umpan balik mengalir secara alami. Jika tugas yang diberikan guru tidak sesuai dengan target pembelajaran maka sulit untuk memberikan umpan balik yang berkualitas.

Latihan

1. Buatlah daftar umpan balik yang Anda biasa berikan kepada siswa! Kategorikan umpan balik tersebut ke dalam strategi-strategi umpan balik yang dibahas dalam bab ini!
2. Tulislah alasan mengapa Anda lebih sering memberikan umpan balik jenis tertentu!

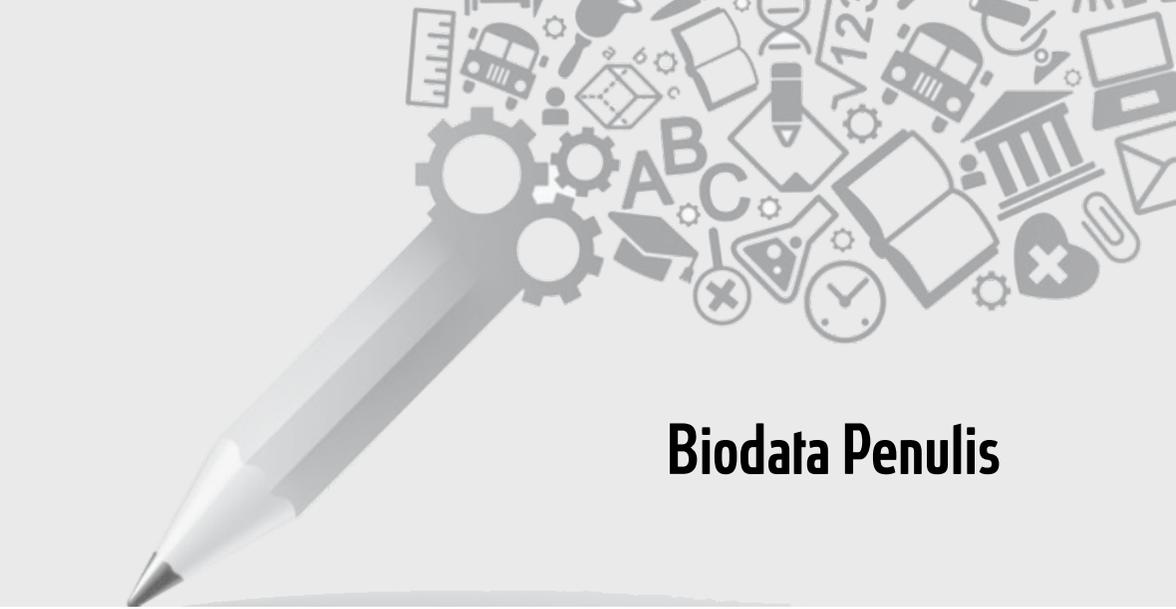
[Halaman ini sengaja dikosongkan]



Daftar Pustaka

- Andrade, H. 2013. "Teaching With Rubrics: The Good, The Bad, And The Ugly". *College Teaching*, 53(1), 27–30.
- Ashraf, H., & Zolfaghari, S. 2018. "EFL Teachers Assessment Literacy And Their Reflective Teaching". *International Journal of Instruction*, 11(1), 425–436.
- Bachman, L. F., & Palmer, a. S. A. S. 1996. "Language Testing in Practice: Designing and Developing Useful Language Tests". *Oxford Applied Linguistics*.
- Brookhart, S. M. 1999. "Teaching About Communicating Assessment Results And Grading". *Educational Measurement: Issues and Practice*, 18(1), 5–13.
- Brookhart, S. M. 2008. *How To Give Effective Feedback To Your Students*. Alexandria, VA: ASCD.
- Brown, H. D. 2004. *Language Assessment: Principles And Classroom Practices*. White Plains, NY: Pearson Education.
- Coombe, C., Al-Hamly, M., & Troudi, S. 2009. "Foreign And Second Language Teacher Assessment Literacy: Issues, Challenges And Recommendations". *Research Notes*, 38(3), 14–18.
- DeLuca, C., & Klinger, D. A. 2010. "Assessment Literacy Development: Identifying Gaps In Teacher Candidates Learning". *Assessment in Education: Principles, Policy & Practice*, 17(4), 419–438.

- Djoub, Z. 2017. Assessment Literacy: Beyond Teacher Practice. In R. Al-Mahrooqi, C. Coombe, F. Al-Maamari, & V. Thakur (Eds.), *Revisiting EFL Assessment: Critical Perspective*. Switzerland: Springer International Publishing.
- Earl, L. 2003. Using Assessment to Motivate Learning. In *Assessment as Learning: Using Classroom Assessment To Maximize Student Learning*. Thousand Oaks, CA: Corwin Press.
- Gronlund, N. E. 1998. *Assessment of Student Achievement*. Boston: Allyn and Bacon Huba.
- Hughes, A. 1993. *Testing For Language Teachers*. London: Cambridge University Press.
- Jannati, S. 2015. "ELT Teachers' Language Assessment Literacy: Perceptions And Practices". *Educational Research Association The International Journal of Research in Teacher Education*, 6(2), 26–37.
- Koh, K. H. 2011. "Improving Teachers Assessment Literacy Through Professional Development". *Teaching Education*, 22(3), 255–276.
- Mcnamara, T. 2009. Validity in Language Testing: The Challenge of Sam Messick's Legacy. *Validity in Language Testing: The Challenge of Sam Messick's Legacy*, (July 2013), 37–41.
- Odo, D. M. 2015. An Online Journal For Improving Urban Teachers Assessment Literacy Through Synergistic Individualized Tutoring And Self-Reflection, 17(2), 1–13.
- Rahman, M. 2018. Teachers Perceptions and Practices of Classroom Assessment in Secondary School Science Classes in, 7(6), 254–263.
- Stiggins, R. J. 1991. "Relevant Classroom Assessment Training For Teachers". *Educational Measurement: Issues and Practice*, 10(1), 7–12.
- Yamtim, V., & Wongwanich, S. 2014. "A Study of Classroom Assessment Literacy Of Primary School Teachers". *Procedia-Social and Behavioral Sciences*, 116, 2998–3004.



Biodata Penulis



Siti Zulaiha, MA., Ph.D. adalah dosen di program studi Magister Pendidikan Bahasa Inggris, Sekolah Pascasarjana (SPs) Universitas Muhammadiyah Prof. DR. HAMKA (UHAMKA) Jakarta. Ia menyelesaikan studi Magister dan Doktor di University of Queensland, Australia. Selain mengajar, ia adalah konsultan pendidikan dan berpengalaman bekerja di berbagai program pendidikan di Indonesia, Australia, dan Filipina. Ia melakukan pelatihan untuk guru-guru dan kepala sekolah serta menulis berbagai modul yang digunakan untuk menunjang pembelajaran di sekolah-sekolah.



Herri Mulyono, Ph.D., lahir di Jakarta, 5 Oktober 1980, adalah dosen di Jurusan Pendidikan Bahasa Inggris, Fakultas Keguruan dan Ilmu Pendidikan (FKIP) Universitas Muhammadiyah Prof. DR. HAMKA (UHAMKA). Selain itu, Beliau juga tercatat sebagai dosen pengampu mata kuliah Qualitative Research in ELT dan ICT in TESOL di Sekolah Pascasarjana (SPs) UHAMKA. Herri menyelesaikan pendidikan sarjana strata satu (S1) di jurusan Pendidikan Bahasa Inggris di Universitas Negeri Jakarta dan pendidikan

sarjana strata dua (S2) di Program Magister Pendidikan Bahasa Inggris SPs UHAMKA, Jakarta. Gelar Doctor of Philosophy (Ph.D) in TESOL diraih dari University of York, Inggris. Selain aktif dalam menulis dan publikasi, Herri aktif sebagai pembicara dalam seminar dan konferensi di dalam maupun luar negeri.