

BAB 4

PENGUJIAN

Dalam bab ini dijelaskan proses pengujian dan perbaikan skema penilaian kelas *TE* tingkat dasar di ILP secara terperinci. Sesuai dengan yang telah dijelaskan dalam bab sebelumnya, pengujian skema penilaian dalam proyek ini dilakukan dengan menggunakan langkah-langkah yang ditawarkan oleh Underhill (1987). Pengujian tersebut diawali dengan penjelasan latar belakang pendidikan siswa dan program yang mereka ikuti, dilanjutkan dengan penjelasan tentang alasan pembuatan skema penilaian dan aspek penilaian seperti deskripsi dan area performansi. Setelah itu dilakukan penilaian bersama, dan pembahasan hasil penilaian awal. Tahap terakhir adalah mengujikannya sekali lagi pada para pengajar dengan menggunakan video performansi yang berbeda. Setelah didapat hasil penilaian akhir, maka penulis melakukan perbaikan skema penilaian.

4.1 Penjelasan Latar Belakang Siswa dan Program

Penjelasan latar belakang siswa dilakukan agar penilai memahami latar pendidikan siswa yang sedang mengikuti program *TE* tingkat dasar. Penilai perlu mengetahui latar pendidikan siswa agar mereka mendapatkan gambaran awal performansi siswa yang akan mereka nilai. Penjelasan latar belakang dilakukan agar penilai dapat memahami tujuan utama diadakannya program ini dan penilaian di dalam program ini.

Ada dua orang siswa yang dijadikan sebagai objek penilaian. Siswa pertama yang digunakan sebagai objek penilaian bernama Novi. Siswa ini saat ini masih berstatus sebagai mahasiswa di sebuah universitas swasta di Jakarta. Menurut pengajarnya, Novi adalah siswa yang sangat aktif dan tidak pemalu untuk mencoba sesuatu yang baru. Sebelumnya, Novi mengikuti kelas *TE* 1. Kelas *TE* ini merupakan kelas *TE* kedua yang diikutinya. Berdasarkan fakta tersebut, penulis dapat menyimpulkan bahwa siswa ini sudah mempelajari bahasa Inggris

sebelumnya. Selain itu, sebagaimana telah disebutkan di dalam bab sebelumnya, siswa yang dapat memasuki kelas *TE* tingkat dasar ini merupakan siswa yang paling tidak sudah menyelesaikan tingkat dasar kelas *GE*. Performansi yang dilakukan oleh siswa ini dijadikan sebagai objek penilaian awal.

Siswa lainnya bernama Erbit. Siswa ini sangat aktif dan cenderung dominan dalam berinteraksi dengan siswa lain. Siswa ini bekerja di sebuah perusahaan swasta di Jakarta. Kelas ini juga merupakan kelas kedua yang diikutinya. Latar belakang kedua orang itu kemudian dijelaskan di bagian awal pelatihan. Setelah penjelasan tersebut, penulis menjelaskan tujuan program dan penilaian. Penulis menjelaskan bahwa program ini dibuat untuk memfasilitasi siswa untuk melatih kemampuan berbicaranya dalam bahasa Inggris. Selain itu, penulis menjelaskan bahwa penilaian di kelas ini dilakukan untuk memberikan balikan kepada siswanya.

Pada tahap pengenalan latar belakang siswa dan program tersebut, pengajar kelas yang bersangkutan menyatakan bahwa kelas ini merupakan kelas yang sangat aktif, dalam arti, semua siswa di kelas ini sangat berusaha untuk berinteraksi dengan sesamanya. Pemberitahuan ini sangat membantu para penilai, karena dengan demikian mereka sudah memiliki pengetahuan akan situasi kelas siswa yang akan dinilainya dan dengan demikian diharapkan dapat mengantisipasi hal-hal yang dapat mengganggu penilaian nantinya.

4.2 Penjelasan Rasional

Pada tahap ini, penulis menjelaskan kepada para penilai yang juga merupakan rekan kerja penulis, alasan pembuatan skema penilaian ini. Penulis menjelaskan bahwa skema penilaian yang sekarang dipakai sangat subjektif. Sebagian besar penilai yang juga pernah mengajar kelas *TE*, sependapat dengan pemikiran penulis. Selain itu, penulis juga menjelaskan bahwa skema penilaian yang sedang dibuat ini dapat digunakan sebagai pengganti skema yang sekarang berlaku dan juga sebagai pendukung proses pemberian balikan.

Pada tahap ini seorang pengajar juga menyarankan bahwa skema penilaian baru ini dapat juga digunakan sebagai tambahan skema yang sudah ada, terutama

dalam hal pemberian angka. Selain itu, seorang pengajar lagi menyatakan bahwa penggunaan skema ini secara berkelanjutan akan menggunakan terlalu banyak kertas. Hal ini merupakan salah satu pertimbangan yang sudah dipikirkan oleh penulis. Jalan keluar yang disarankan oleh penulis adalah dengan menggunakan pensil untuk memberi penilaian dan menyediakan pengarsipan khusus untuk menyimpan penilaian yang telah dilakukan per hari. Semua pengajar di dalam pelatihan sependapat dengan hal tersebut. Selanjutnya, setelah diarsipkan, skema penilaian yang telah diisi dapat dihapus dan digunakan kembali untuk hari lainnya. Seorang pengajar menyarankan untuk memberikan nama di bagian atas skema penilaian sehingga skema yang telah dihapus dapat digunakan untuk siswa yang sama. Saran tersebut disetujui oleh penulis dan pengajar lainnya.

4.3 Penjelasan Aspek Penilaian

Penjelasan aspek penilaian dilakukan dengan pengisian data kuesioner (lihat lampiran) yang dikorelasikan dengan skema penilaian. Kuesioner tersebut berisi kriteria keterampilan bahasa komunikatif yang dikemukakan oleh Bachman (1990) sebagaimana ditafsirkan oleh Qin and Li (2008). Kuesioner ini dibuat untuk mendapatkan pengertian pengajar tentang konsep kompetensi dan untuk memperkenalkan mereka kepada konsep tersebut jika mereka belum mengenalinya. Selain itu, kuesioner ini ditujukan untuk memeriksa apakah pengajar setuju untuk memasukkan kriteria kompetensi ke dalam suatu skema penilaian. Seorang pengajar banyak bertanya tentang kuesioner dan skema penilaian karena ia tidak mengenal sebagian istilah yang dituliskan di dalam skema tersebut. Sebagai tambahan, pelatihan dan pengujian ini untuk sementara dilakukan terhadap lima (5) orang pengajar yang sebagian besar sudah memiliki pengalaman mengajar di kelas *TE* selama 1 tahun sampai dengan 16 tahun. Penulis berjanji akan melakukan pelatihan lebih lanjut terhadap pengajar dengan jumlah yang lebih banyak untuk memaksimalkan keandalan dan kesahihan skema penilaian. Para pengajar tersebut diharuskan untuk melingkari satu pilihan jawaban dari empat yang tersedia. Jawaban bernomor 1 mewakili pernyataan sangat tidak setuju sementara no 4

mewakili pernyataan sangat setuju (lihat Lampiran 4). Berikut hasil kuesioner beserta analisisnya.

Penafsiran Simpangan Baku:

0.0 – 1.0 = keandalan kriteria lemah

1.0 – 2.0 = keandalan kuat

2.0 – 3.0 = keandalan sangat kuat

Tabel 4.1.
Analisis Hasil Kuesioner

Nomor Item	1	2	3	4	Rata-rata	Simpangan Baku
1	0 (0%)	1 (20%)	4 (80%)	0 (0%)	3.5	1.27
2	0 (0%)	1 (20%)	4 (80%)	0 (0%)	3.5	1.27
3	0 (0%)	2 (40%)	3 (60%)	0 (0%)	3.25	0.6
4	0 (0%)	0 (0%)	4 (80%)	1 (20%)	4	1.5
5	0 (0%)	0 (0%)	4 (80%)	1 (20%)	4	1.5
6	0 (0%)	1 (20%)	1 (20%)	3 (60%)	4.25	2.3
7	0 (0%)	0 (0%)	4 (80%)	1 (20%)	4	1.5
8	0 (0%)	0 (0%)	2 (40%)	3 (60%)	4.5	1.45
9	0 (0%)	0 (0%)	4 (80%)	1 (20%)	4	1.5
10	0 (0%)	0 (0%)	1 (20%)	4 (80%)	4.75	1.9
11	0 (0%)	0 (0%)	3 (60%)	2 (40%)	4.25	1.2
12	0 (0%)	1 (20%)	4 (80%)	0 (0%)	3.5	1.27

Item 1 sampai dengan 3 (Lampiran 4) mengacu kepada kompetensi gramatikal atau yang di dalam skema ini disebut sebagai pengetahuan gramatikal. Item 1 mengacu pada kemampuan menyusun kata menjadi suatu kalimat atau ujaran dengan struktur yang baik. Item 2 berkaitan dengan pelafalan yang baik sementara item 3 mengacu pada penguasaan kosa kata. Seperti yang terlihat di dalam tabel berikut, para pengajar sepakat bahwa komponen pengetahuan gramatikal merupakan salah satu komponen utama kompetensi berbicara dengan nilai rata-rata di atas 3 dengan simpangan baku di atas angka 1 yaitu 1.27. Dari

tabel ini, dapat disimpulkan pula bahwa penguasaan tata bahasa dianggap tidak terlalu mewakili kompetensi berbicara. Ini terlihat dari nilai rata-rata yang rendah dengan simpangan baku di bawah 1 yaitu 0.6. Hal ini sangat masuk akal karena penutur jati pun kadang berujar tanpa menggunakan struktur bahasa yang sesuai dengan kaidah.

Item 4 dan 5 (Lampiran 4) mewakili kompetensi organisasional. Item 4 mengacu pada kompetensi penggunaan kata sambung sementara item 5 erat kaitannya dengan kemampuan siswa menyusun ujarannya dalam suatu kondisi atau situasi. Para pengajar sepakat untuk memasukkan kompetensi ini ke dalam skema penilaian. Hal tersebut bisa dilihat dari rata-rata item tersebut yang bernilai 4 dengan simpangan baku 1.5. Terlebih lagi, dari penilaian tersebut, penulis menyimpulkan bahwa kemampuan menyusun ujaran dalam suatu wacana secara rapi dan padu lebih penting dari sekedar akurasi ujaran.

Item 6 (Lampiran 4) terkait dengan kompetensi ilokusioner. Kriteria yang ditanyakan kepada para pengajar adalah kriteria penggunaan bahasa. Dari tabel tersebut dapat kita lihat bahwa kemampuan siswa untuk dapat menggunakan bahasa yang dipelajarinya dalam situasi dan kondisi tertentu sangat dianggap penting. Ini dibuktikan dengan akumulasi rata-rata kriteria tersebut yang memiliki nilai rata-rata 4.25 dengan simpangan baku di atas angka 1 yaitu 2.3.

Item 7 dan 8 (Lampiran 4) merupakan perwakilan dari area performansi kompetensi sociolinguistik. Item 7 terkait dengan penggunaan laras dan gaya bahasa dalam ujaran, sementara item 8 mengacu pada kealamian berbahasa. Kedua kriteria ini juga mendapatkan rata-rata tinggi dengan nilai di atas 4 (4 dan 4.5 dengan simpangan baku 1.5 dan 1.45). Dari fakta tersebut, penulis dapat menyimpulkan bahwa kealamian berbahasa masih merupakan ukuran kefasihan dan kompetensi seseorang dalam berujar.

Komponen kompetensi strategis diwakilkan di item 9 sampai dengan 11 (Lampiran 4). Item 9 berkaitan dengan analisis dan pemilihan tugas. Item 10 mewakili kemampuan siswa dalam menanggapi ujaran orang lain sementara item 11 dan 12 erat adalah tentang kemampuan siswa untuk menyampaikan pesan dengan

menggunakan berbagai cara, salah satunya adalah penggunaan bahasa tubuh. Dari tabel di atas dapat kita lihat penggunaan bahasa tubuh tidak terlalu dianggap sebagai salah satu kemampuan yang harus dimiliki siswa. Ini terlihat dari rata-rata nilai yang diperolehnya yaitu 3.5 dengan simpangan baku 1.27. Sementara itu, komponen kompetensi strategis yang dianggap paling penting oleh para pengajar adalah kemampuan siswa dalam menanggapi ujaran lawan bicaranya. Nilai rata-rata untuk kriteria ini adalah 4.75 dengan simpangan baku 1.9, yang juga merupakan nilai rata-rata tertinggi dari semua kriteria performansi. Hal ini sejalan dengan pemikiran Bachman (1990) yang memang mengutamakan kompetensi strategis di dalam model keterampilan bahasa komunikatif yang dibuatnya.

Berdasarkan analisis tersebut, penulis dapat menyimpulkan bahwa para pengajar sepakat untuk menggunakan kelima komponen kompetensi tersebut sebagai bahan pertimbangan penilaian performansi. Selain itu, dengan pengisian kuesioner ini, para pengajar juga dapat memahami konsep kompetensi yang dimiliki oleh seseorang.

4.4. Penilaian Awal

Tujuan utama penilaian awal adalah membiasakan penilai untuk menggunakan skema penilaian beserta deskripsi performansinya. Penilaian dilakukan atas performansi siswa bernama Novi yang sedang menggambarkan rumah yang ingin dimilikinya. Pelajaran menggambarkan keadaan dan ruangan rumah telah dipelajari pada pertemuan sebelumnya. Penilaian tersebut dilakukan atas performansi yang berdurasi kurang-lebih 3 menit sebagaimana terekam di dalam video kamera. Seperti telah disebutkan di dalam subbab terdahulu, siswa yang dinilai di kelas ini sangat aktif jadi terdapat sedikit kesulitan dalam penilaiannya dikarenakan keadaan kelas yang sangat ramai. Oleh karena itu, penulis harus mengulang rekaman video sebanyak satu kali. Setelah diputar sebanyak dua kali, penilaian yang dilakukan oleh kelima pengajar tersebut dapat terlihat di dalam tabel berikut. Kriteria penilaian di dalam tabel berikut akan disingkat untuk mempermudah pembuatan rekapitulasi. **PG** merupakan singkatan dari

pengetahuan gramatikal, **O** untuk organisasi, **FB** untuk fungsi bahasa, **K** untuk kewajaran, dan **SB** adalah singkatan untuk strategi berbahasa. Penyingkatan lain juga dilakukan atas penamaan skala, yaitu **Ma** untuk mahir, **Me** untuk memuaskan, dan **Pe** untuk pemula.

Tabel 4.2
Rekapitulasi Penilaian Awal

Kriteria	Pengajar				
	Ibu Shinta	Ibu Ria	Ibu Dina	Ibu Yanna	Bapak Frantze
PG	Me	Me	Me	Me	Me
O	Ma	Me	Me	Ma	Me
FB	Ma	Me	Ma	Ma	Ma
K	Ma	Me	Me	Me	Me
SB	Ma	Ma	Ma	Ma	Ma

Lebih lanjut, seperti disebutkan di bab 3, penafsiran keandalan kualitatif ini disajikan pula dalam bentuk kuantitatif. Untuk penyusunan data kuantitatif rekapitulasi penilaian awal, penulis menyusun data berdasarkan penjelasan Salkind (2004). Ia menyarankan agar keandalan antarpenilai disusun berdasarkan perbandingan jumlah jawaban sama terbanyak dengan jumlah kemungkinan jawaban yang sama.

$$IR = A/TPA$$

Berdasarkan perhitungan tersebut, penulis juga telah menentukan skala keandalan penilaian awal dengan ketentuan seperti yang terlihat di tabel 4.3 berikut. Penyusunan skala itu juga mengacu pada pernyataannya tentang penentuan

keandalan penilaian. Bagan tersebut kemudian digunakan untuk menafsirkan tabel 4.4.

Tabel 4.3
Tabel Keandalan

Nilai	Pernyataan Keandalan
0.8 – 1.0	keandalan sangat kuat
0.6 – 0.8	keandalan kuat
0.4 – 0.6	keandalan lemah
0.2 – 0.4	keandalan sangat lemah
0 – 0.2	tidak ada keandalan

Tabel 4.4
Tabel Keandalan Penilaian Awal

Kriteria	Terbanyak	Total	Presentase	S k a l a Keandalan
PG	5	5	100 %	1
O	3	5	60 %	0.6
FB	4	5	80 %	0.8
K	4	5	80 %	0.8
SB	5	5	100 %	1
Rata-rata Keandalan				0.84

Dari tabel di atas dapat terlihat bahwa kriteria PB dan SB sudah memiliki unsur keandalan antar penilai yang sangat kuat dengan nilai keandalan 1. Hal ini dibuktikan dengan kesamaan penilaian yang diberikan pada kriteria tersebut oleh

lima orang pengajar yang diminta memberikan penilaian atas performa siswa tersebut. Terlebih lagi, skema ini, secara keseluruhan, memiliki keandalan antar penilai dengan nilai total keandalan 0.84.

Pertimbangan unsur kesahihan juga telah terbukti dalam tabel ini. Seperti telah disebutkan terdahulu, ada dua macam kesahihan yang akan dipertimbangkan dalam pembuatan skema ini yaitu kesahihan isi dan muka. Kesahihan isi dibuktikan dengan isi kriteria penilaian yang memang sejalan dengan pembuatan program, yaitu pemberian balikan maksimal dengan tidak mengganggu kesempatan siswa untuk mengekspresikan diri (lihat lampiran 1). Kesahihan muka dibuktikan dengan pertanyaan yang ditanyakan setelah penilaian dengan cara penulis menanyakan apakah skema ini dapat digunakan untuk penilaian berkelanjutan yang kemudian disepakati oleh para pengajar. Hal ini sejalan dengan pernyataan Brown (2004) tentang pengukuran kesahihan muka.

Unsur praktis juga telah terpenuhi di dalam skema ini. Brown (2004) menyebutkan bahwa suatu penilaian praktis jika memenuhi persyaratan tertentu, yaitu tidak mahal, efisien dan mudah digunakan. Penyusunan skema ini hanya membutuhkan dua buah lembar kertas yang bisa digunakan sepanjang periode pembelajaran. Selain itu, karena penilaian ini menggunakan sistem *check list*, skema ini sangat efisien dan tidak memakan waktu. Prosedur penilaian pun sangat efisien dan tidak rumit, ini dibuktikan dengan tidak adanya pertanyaan ketika penulis menelaskkan prosedur penilaian. Penulis juga cukup hanya menjelaskan prosedur tersebut satu kali. Itu merupakan bukti bahwa skema ini sangat praktis untuk digunakan.

4.5. Pembahasan Hasil Penilaian

Meskipun terdapat persamaan, perbedaan penilaian masih mendominasi penilaian awal dengan menggunakan skema tersebut. Seperti yang terlihat di tabel 4.2., untuk kriteria organisasi, dua orang memberikan penilaian berbeda. Ketika dibahas perbedaan tersebut, diketahui bahwa kedua penilai tersebut memiliki alasan yang serupa dalam pemberian nilai. Ibu Shinta dan ibu Yanna menyatakan bahwa

pemberian "Ma" kepada siswa yang bersangkutan dikarenakan kedua pengajar tersebut sangat mendetil dalam pemberian nilainya. Keduanya menyatakan bahwa Novi telah menggunakan menggunakan kata sambung yang memadai, sementara ketiga orang lainnya tidak memperhatikan hal tersebut. Oleh karena itu, penulis kembali mengingatkan para penilai untuk kembali membaca dengan teliti kriteria yang terdapat dalam setiap area performansi.

Untuk kriteria FB, hanya satu orang penilai yang memberikan penilaian yang berbeda, yaitu ibu Ria. Ketika ditanyakan penyebab perbedaan tersebut, ia menyatakan bahwa ia menyadari telah memperhatikan kriteria yang salah dalam penilaian tersebut. Ia menyebutkan bahwa siswa terlalu lama berpikir ketika membuat ujaran. Ibu Ria tidak menyadari bahwa kategori itu telah dipertimbangkan di dalam penilaian untuk kriteria PG. Setelah diberi penjelasan, tidak ada pertanyaan lagi tentang perbedaan penilaian untuk kriteria FB.

Perbedaan penilaian untuk kriteria kewajaran lebih disebabkan oleh preferensi pengajar dalam memberikan nilai. Ibu Shinta menyebutkan bahwa ia sangat menghargai muridnya, sehingga memberikan nilai "Ma" untuknya. Itu membuktikan bahwa tanpa pelatihan yang benar, subjektivitas penilaian sangat mungkin, bahkan pada penilaian yang seharusnya bersifat objektif. Para pengajar dan penulis kemudian menyarankan kepada Ibu Shinta bahwa untuk memberikan balikan yang maksimal, seorang penilai harus bersifat objektif. Ibu Shinta kemudian sepakat untuk lebih objektif dalam penilaiannya untuk performansi berikutnya.

4.6. Penilaian Kedua

Penilaian keandalan skema atau penilaian kedua dilakukan setelah para penilai terbiasa dengan penggunaan skema penilaian dan pembahasan kriteria penilaian. Penilaian kedua dilakukan terhadap siswa yang berbeda yaitu seorang siswa bernama Erbit. Pemilihan objek penilaian yang berbeda dilakukan berdasarkan beberapa alasan. Alasan pertama adalah untuk menghindari terjadinya penguatan performa yang kadang terjadi jika seseorang melakukan penilaian terhadap objek yang sama. Selain itu, pemilihan ini juga ditujukan untuk

memastikan unsur keandalan. Penilaian dilakukan atas performansi siswa yang direkam secara audio visual berdurasi lebih-kurang 3 menit. Namun, tidak seperti pada rekaman pertama, suasana di kelas tidak begitu gaduh sehingga penulis hanya memutar rekaman sebanyak satu kali. Rekapitulasi penilaian atas performansi kedua itu dapat dilihat pada tabel berikut. Di bawah tabel rekapitulasi tersebut dapat pula dilihat tabel 4.6 tentang keandalan untuk penilaian skema.

Tabel 4.5
Rekapitulasi Penilaian Keandalan Skema Penilaian

Kriteria	Pengajar				
	Ibu Shinta	Ibu Ria	Ibu Dina	Ibu Yanna	Bpk.Frantze
PG	Me	Ma	Ma	Ma	Ma
O	Ma	Ma	Ma	Ma	Ma
FB	Ma	Ma	Ma	Ma	Ma
K	Ma	Ma	Ma	Ma	Ma
SB	Ma	Ma	Ma	Ma	Ma

Tabel 4.6
Tabel Keandalan Penilaian Kedua

Kriteria	Terbanyak	Total	Presentase	S k a l a Keandalan
PG	4	5	80 %	0.8
O	5	5	100 %	1
FB	5	5	100 %	1
K	5	5	100 %	1

SB	5	5	100 %	1
Rata-rata Keandalan				0.96

Dari tabel tersebut terlihat bahwa skema penilaian ini memiliki tingkat keandalan yang sangat tinggi karena terdapat persamaan penilaian untuk kriteria organisasi sampai dengan strategi berbahasa. Hanya terdapat satu perbedaan yang diberikan oleh ibu Shinta untuk kriteria pengetahuan gramatikal. Ketika ditanyakan penyebabnya, ibu Shinta mengaku lupa bahwa kesalahan gramatikal diperbolehkan bahkan untuk skala mahir. Dengan penjelasan itu, ia pun mengubah penilaiannya.

4.7. Tindak Lanjut

Berdasarkan pengujian skema penilaian terdapat beberapa temuan yang perlu ditindaklanjuti. Temuan pertama berkaitan dengan jumlah responden yang digunakan untuk pengujian. Seperti terlihat di dalam tabel 4.3 tentang pengujian skema penilaian, pengujian sesungguhnya atas performansi siswa menunjukkan hasil yang memuaskan. Kesalahan yang terjadi banyak disebabkan oleh faktor perseorangan. Meskipun demikian, perlu diingat bahwa pengujian itu dilakukan hanya dengan lima orang peserta. Oleh karena itu, pengujian lebih lanjut akan dilakukan terhadap responden yang lebih banyak untuk mempertegas keandalan yang telah diraih skema penilaian ini.

Temuan lain berkaitan dengan pelatihan skema penilaian, terutama tentang permasalahan subjektifitas. Seperti yang tersebut di dalam bab terdahulu, beberapa pengajar masih sering bersifat subjektif di dalam penilaiannya. Oleh karena itu, jika skema ini digunakan oleh semua pengajar *TE*, perlu ditegaskan pentingnya untuk bersikap objektif karena objektivitas dalam penilaian dapat menghasilkan balikan yang maksimal. Pemberian balikan yang maksimal, selain berguna untuk penilaian, juga sangat berguna bagi perkembangan kompetensi siswa.

Terkait dengan konversi skema penilaian ke dalam angka, penulis ini telah memutuskan untuk tidak melakukannya karena, sejak awal, pembuatan skema ini