

BAB III

METODE PENELITIAN

3.1 Pendekatan Penelitian

Penelitian ini merupakan penelitian evaluasi dengan menggunakan pendekatan deskriptif kuantitatif dan model evaluasi CIPP (*Context, Input, Process, Product*) yang dikembangkan oleh Stufflebeam. Model ini dipilih karena merupakan model evaluasi yang sangat tepat untuk mengevaluasi efektivitas pelaksanaan sebuah program, termasuk evaluasi program. Model CIPP ini, terdiri dari empat komponen, yakni:

1. Context evaluation

Evaluasi konteks adalah upaya untuk menggambarkan kondisi awal. Dalam penelitian ini, evaluasi konteks diarahkan pada visi misi, pengelolaan, kepemimpinan, dan sistem informasi manajemen program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* Bandar Jaya.

2. Input evaluation

Evaluasi masukan ditujukan pada ketersediaan fasilitas sarana prasarana, sumber daya manusia dan kurikulum program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* Bandar Jaya.

3. *Process evaluation*

Evaluasi proses diarahkan pada seberapa jauh kegiatan yang dilaksanakan di dalam program yang sudah terlaksana sesuai rencana. Dalam penelitian ini, evaluasi proses berfokus pada perencanaan, pelaksanaan dan evaluasi program pembelajaran Bahasa Inggris kelas *English for Children*.

4. *Product evaluation*

Evaluasi produk diarahkan pada pencapaian hasil belajar program pembelajaran Bahasa Inggris kelas *English for Children* setelah mengikuti program pembelajaran Bahasa Inggris.

3.2 Tempat dan Waktu Penelitian

1. Tempat

Tempat yang digunakan untuk penelitian adalah lembaga kursus *English Smart* Bandarjaya dengan alamat Jl. Imam Bonjol No.3 Yukum Jaya, Bandar Jaya kec. Terbanggi Besar kab. Lampung Tengah. Pemilihan *English Smart* Bandar Jaya sebagai tempat penelitian ini karena lembaga kursus tersebut merupakan salah satu lembaga kursus Bahasa Inggris yang terbesar di Lampung Tengah dan pembelajaran oleh tentor belum pernah dilakukan evaluasi.

2. Waktu

Pelaksanaan penelitian ini pada satu periode level yaitu bulan Juli sampai Oktober 2014.

3.3 Subjek Penelitian

Subjek penelitian dalam penelitian ini meliputi: kepala lembaga, supervisor, tentor sebanyak 5 orang dan pembelajar dari 5 kelas *English for Children* berjumlah 48 orang yang berada di *English Smart* Bandarjaya.

Tabel 3.1 Subjek Penelitian

No	Subjek Penelitian	Jumlah
1.	Kepala Lembaga <i>English Smart</i>	1
2.	<i>Supervisor</i>	1
3.	Tentor kelas <i>English for Children</i>	5
4.	Pembelajar kelas <i>English for Children</i>	48
Jumlah Total		55

Sumber : Staf Tata Usaha *English Smart* Bandar Jaya

Dasar pengambilan subyek dalam penelitian ini berpedoman pada pendapat Arikunto (2010:112), bahwa apabila jumlah subjeknya kurang dari 100, lebih baik diambil semua sehingga penelitiannya merupakan penelitian populasi. Jika jumlah subjeknya besar dapat diambil antara 10% - 15% atau 20% - 25% atau lebih. Dari data tersebut dapat dijelaskan populasi dalam penelitian ini sebesar 55 orang.

3.4 Teknik Pengumpulan Data

Pengumpulan data adalah prosedur yang sistematis dan standar untuk memperoleh data yang diperlukan. Beberapa cara untuk mengumpulkan data yaitu:

1. Observasi

Metode observasi atau yang dikenal juga dengan pengamatan. Instrumen yang digunakan adalah pedoman observasi dan *check list*. Subyek penelitian adalah

kepala lembaga, *supervisor*, dan tentor. Tujuan menggunakan metode ini yaitu mengumpulkan data komponen:

- a. *Context*, tentang kondisi awal
- b. *Input*, tentang ketersediaan sarana prasarana, tenaga pendidik dan kurikulum lembaga.
- c. *Process*, tentang perencanaan, pelaksanaan dan evaluasi program pembelajaran Bahasa Inggris.

2. Dokumentasi

Dokumentasi digunakan untuk mengumpulkan data *product*, mencari data mengenai variabel yang berupa catatan, transkrip, buku, agenda dan sebagainya. Dengan metode dokumentasi ini peneliti akan memperoleh data sekunder tentang komponen *context*, *input*, *process* dan *product*.

3. Tes

Tes digunakan untuk mengumpulkan data *product* yang berupa tes akhir pembelajar yang diberikan setelah selesai pembelajaran Bahasa Inggris pada satu periode belajar.

3.5 Definisi Konseptual dan Operasional

1. Definisi Konseptual

Evaluasi konteks merupakan dasar dari evaluasi yang bertujuan menyediakan alasan-alasan dalam penentuan tujuan karenanya upaya yang dilakukan evaluator dalam evaluasi konteks ini adalah memberikan gambaran dan rincian terhadap lingkungan, kebutuhan serta tujuan. Evaluasi konteks mencakup analisis masalah yang berkaitan dengan lingkungan program atau kondisi obyektif yang akan dilaksanakan. Evaluasi *context* dalam program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* Bandar Jaya mencakup kondisi awal lembaga yaitu meliputi visi misi, pengelolaan, kepemimpinan, dan sistem informasi manajemen.

Evaluasi input merupakan evaluasi yang bertujuan menyediakan informasi untuk menentukan bagaimana menggunakan sumberdaya yang tersedia dalam mencapai tujuan program. Evaluasi input meliputi analisis personal yang berhubungan dengan bagaimana penggunaan sumber-sumber yang tersedia, alternatif-alternatif strategi yang harus dipertimbangkan untuk mencapai suatu program. Evaluasi *input* dalam program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* Bandar Jaya meliputi ketersediaan sarana prasarana, tenaga pendidik dan kurikulum yang digunakan.

Evaluasi yang dirancang dan diaplikasikan dalam praktik implementasi kegiatan disebut dengan evaluasi proses. Untuk melihat apakah pelaksanaan program sudah sesuai dengan strategi yang telah dilaksanakan tersebut, maka perlu diadakannya

evauasi. Evaluasi *process* dalam program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* Bandar Jaya meliputi perencanaan, pelaksanaan dan evaluasi pembelajaran

Evaluasi produk adalah evaluasi mengukur keberhasilan pencapaian tujuan. Evaluasi dapat juga bertujuan mengumpulkan deskripsi dan penilaian terhadap iuran (*outcome*) dan menghubungkan itu semua dengan objektif, konteks, input, dan proses serta untuk menginterpretasikan kelayakan dan keberhargaan program. Evaluasi *product* dalam program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* Bandar Jaya meliputi hasil belajar pembelajar.

2. Definisi Operasional

Sesuai dengan model evaluasi yang digunakan dalam penelitian ini yaitu model evaluasi CIPP, maka definisi variabel secara operasional adalah sebagai berikut:

a. Evaluasi Konteks (*Context*)

Evaluasi konteks dalam program pembelajaran Bahasa Inggris menggambarkan tentang kondisi awal lembaga yaitu berhubungan meliputi visi misi, pengelolaan, kepemimpinan, dan sistem informasi manajemen yang diperoleh dari data hasil observasi.

b. Evaluasi *Input*

Evaluasi input dalam program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* Bandar Jaya meliputi, ketersediaan sarana prasarana

yang menunjang, sumber daya manusia dan kurikulum. Ketersediaan sarana dan prasarana yaitu meliputi ruang (belajar dan staff) dan ketersediaan buku penunjang pembelajaran. Sumber daya manusia meliputi kualifikasi pendidikan setiap tutor. Sedangkan kurikulum meliputi acuan penggunaan dan peninjauan kurikulum yang diperoleh dari data hasil observasi.

c. Evaluasi Proses (*Process*)

Evaluasi proses dalam program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* menggambarkan tentang proses perencanaan pembelajaran Bahasa Inggris yang meliputi bagaimana tutor mempersiapkan pembelajaran dari penyusunan silabus hingga pembuatan *lesson plan* dan bahan ajar. Proses pelaksanaan pembelajaran Bahasa Inggris yang terlihat di ruang kelas pada saat tutor mengajar yang diperoleh dari data hasil observasi peneliti di *English Smart* Bandar Jaya. Serta evaluasi pelaksanaan pembelajaran yang berfokus pada penilaian saat dan setelah satu periode belajar dilaksanakan.

d. Evaluasi Produk (*Product*)

Evaluasi produk dalam program pembelajaran Bahasa Inggris kelas *English for Children* di *English Smart* Bandar Jaya adalah penilaian hasil belajar selama satu periode belajar. Hasil belajar siswa diperoleh melalui tes.

3.6 Teknik Analisis Data

Data penelitian diperoleh dari hasil evaluasi terhadap setiap variabel penelitian berdasarkan indikator penelitian yang disusun menjadi instrumen penelitian. Untuk menentukan data penelitiannya maka dalam evaluasi diperlukan kriteria penilaian untuk setiap instrumen. Kriteria itu berdasarkan kriteria empiris yaitu kriteria yang dikembangkan di lapangan dengan kriteria kuantitatif dan kualitatif. Masing-masing jenis kriteria dengan pertimbangan dan tanpa pertimbangan. Widoyoko (2012:113) berpendapat bahwa keduanya tetap ilmiah karena disusun berdasarkan penalaran yang benar.

Skoring dilakukan dengan skala bertingkat dengan model skala lima (Arikunto, 2010:8) yaitu terdapat lima pilihan yang disusun dalam bentuk suatu pernyataan dan diikuti oleh pilihan respon yang menunjukkan tingkatan. Kriteria evaluasi keseluruhan tersebut dapat dihitung sebagai berikut :

$$\text{Nilai Kriteria} = \frac{\text{skor yang diperoleh}}{\text{skor maksimal}} \times 100$$

Sebagai contoh yang diperoleh evaluasi program adalah 25 maka nilai kriteria sebagai berikut :

$$\text{Nilai Kriteria} = \frac{25}{45} \times 100$$

Maka nilai kriteria sebesar 55,5 dikategorikan pada hasil kriteria yang cukup.

1. Baik sekali, jika mencapai 81-100%
2. Baik, jika mencapai 61-80%
3. Cukup, jika mencapai 41-60%
4. Kurang, jika mencapai 21-40%
5. Kurang sekali, jika mencapai <21%

3.7 Kriteria Evaluasi

Tabel 3.2 Kriteria Evaluasi

Komponen	Sub Komponen	Indikator	Kriteria Evaluasi	Kriteria Penilaian
<i>Context</i>	Kondisi awal	<p><i>Visi Misi</i></p> <ul style="list-style-type: none"> • Lembaga penyelenggara Program Kursus Bahasa Inggris seharusnya merumuskan dan menetapkan visi, misi, dan tujuan serta memiliki dokumennya • Lembaga penyelenggara Program Kursus Bahasa Inggris sebaiknya melaksanakan sosialisasi visi, misi dan tujuan kepada semua pendidik, peserta didik, dan unsur lain yang terkait <p><i>Pengelolaan</i></p> <ul style="list-style-type: none"> • Lembaga penyelenggara Program Kursus Bahasa Inggris sebaiknya mempunyai pedoman yang mengatur berbagai aspek pengelolaan secara tertulis yang mudah dibaca oleh pihak terkait yang meliputi; Kurikulum, Kalender Pendidikan, Peraturan • Pelaksanaan Program Kursus Bahasa Inggris seharusnya berdasarkan rencana kerja tahunan yang telah ditetapkan <p><i>Kepemimpinan</i></p> <ul style="list-style-type: none"> • Pimpinan Lembaga Penyelenggara Program Kursus Bahasa Inggris harus memiliki kompetensi mengelola serta pengetahuan tentang Bahasa Inggris <p><i>Sistem Informasi Manajemen</i></p> <ul style="list-style-type: none"> • Lembaga Penyelenggara Program Kursus Bahasa Inggris sebaiknya 	Sesuai dengan Standar Nasional Pendidikan (SNP) Peraturan Pemerintah Republik Indonesia Nomor 19 tahun 2005 tentang standar pengelolaan lembaga kursus dan pelatihan	Sangat Baik

Komponen	Sub Komponen	Indikator	Kriteria Evaluasi	Kriteria Penilaian
		menyediakan fasilitas informasi yang efisien, efektif dan mudah diakses		
<i>Input</i>	Sarana dan prasarana	<ul style="list-style-type: none"> • Lembaga penyelenggara Program Kursus Bahasa Inggris harus memiliki tempat aktifitas belajar (ruang teori/praktek) • Lembaga penyelenggara Program Kursus Bahasa Inggris sebaiknya memiliki ruang aktifitas yang lain (ruang perpustakaan, ruang pendidik, ruang tata usaha dan ruang pimpinan) • Ruang belajar Program Kursus Bahasa Inggris harus dilengkapi berupa alat dan perlengkapan untuk melakukan praktek (“<i>best practice</i>”) Bahasa Inggris • Lembaga Penyelenggara Program Kursus Bahasa Inggris seharusnya menyediakan modul, bahan ajar, handout, yang diperlukan 	Sesuai dengan Standar Nasional Pendidikan (SNP) Peraturan Pemerintah Republik Indonesia Nomor 19 tahun 2005 tentang standar sarana dan prasarana lembaga kursus dan pelatihan	Sangat Baik

Komponen	Sub Komponen	Indikator	Kriteria Evaluasi	Kriteria Penilaian
	Sumber daya manusia Dokumen Kurikulum	<ul style="list-style-type: none"> • Program Kursus Bahasa Inggris harus memiliki pendidik yang memenuhi kompetensi sesuai dengan standar yang ditetapkan • Pendidik Program Kursus Bahasa Inggris harus mengikuti pelatihan peningkatan mutu yang relevan • Lembaga Penyelenggara Program Kursus Bahasa Inggris seharusnya memiliki kurikulum • Kurikulum yang digunakan harus mengacu kepada Standar yang berlaku • Kurikulum seharusnya ditinjau secara berkala • Frekuensi peninjauan/perubahan kurikulum sebaiknya dilakukan secara tahunan/bulanan 	Sesuai dengan Standar Nasional Pendidikan (SNP) Peraturan Pemerintah Republik Indonesia Nomor 19 tahun 2005 tentang standar pendidik lembaga kursus dan pelatihan Sesuai dengan Standar Nasional Pendidikan (SNP) Peraturan Pemerintah Republik Indonesia Nomor 19 tahun 2005 tentang standar isi lembaga kursus dan pelatihan	
<i>Process</i>	Perencanaan pembelajaran	<i>Kalender Pendidikan</i> <ul style="list-style-type: none"> • Lembaga Penyelenggara Program Kursus Bahasa Inggris seharusnya memiliki kalender pendidikan dan disosialisasikan kepada pihak-pihak yang berkepentingan. <i>Silabus</i> <ul style="list-style-type: none"> • Lembaga Penyelenggara Program Kursus Bahasa Inggris seharusnya 	Sesuai dengan Peraturan Pemerintah Republik Indonesia Nomor 19 tahun 2005 tentang standar proses perencanaan pembelajaran lembaga kursus dan pelatihan	Sangat Baik

Komponen	Sub Komponen	Indikator	Kriteria Evaluasi	Kriteria Penilaian
		<p>memiliki silabus</p> <ul style="list-style-type: none"> • Silabus harus disusun dengan mengacu pada standar kompetensi • Silabus seharusnya disusun oleh pendidik • Silabus sebaiknya didokumentasikan <p><i>Lesson Plan</i></p> <ul style="list-style-type: none"> • Lembaga penyelenggara program harus memiliki atau membuat Lesson Plan • Lesson Plan seharusnya disusun berdasarkan silabus mata pelajaran • Lesson Plan setiap mata pelajaran seharusnya disusun oleh Pendidik <p><i>Beban Belajar</i></p> <ul style="list-style-type: none"> • Beban belajar ditetapkan berdasarkan jumlah jam belajar per satuan waktu <p><i>Bahan ajar</i></p> <ul style="list-style-type: none"> • Bahan ajar sebaiknya ditetapkan oleh lembaga. 		
	Pelaksanaan pembelajaran	<p><i>Metode</i></p> <ul style="list-style-type: none"> • Kegiatan awal, membangkitkan motivasi dan memfokuskan perhatian peserta didik agar siap terlibat aktif dalam proses pembelajaran • Kegiatan inti, proses pembelajaran untuk mencapai KD dilakukan secara interaktif, inspiratif, menyenangkan, menantang, memotivasi peserta didik untuk berpartisipasi aktif. Dilakukan secara sistematis dan sistemik melalui proses eksplorasi, elaborasi, dan konfirmasi 	Sesuai dengan Peraturan Menteri Pendidikan Nasional Republik Indonesia No.41 Tahun 2007 tentang Standar proses standar nasional pendidikan	

Komponen	Sub Komponen	Indikator	Kriteria Evaluasi	Kriteria Penilaian
		<ul style="list-style-type: none"> • Penutup, mengakhiri aktivitas pembelajaran ; rangkuman/kesimpulan, penilaian dan refleksi, umpan balik, perlakuan tindak lanjut. <p><i>Media</i></p> <ul style="list-style-type: none"> • Belajar dan pembelajaran, meliputi elektronik dan nonelektronik sesuai dengan SK dan KD, serta Standar Sarana Prasarana 		
	Evaluasi pembelajaran	<p><i>Penilaian</i></p> <ul style="list-style-type: none"> • Penyelenggara program kursus Bahasa Inggris seharusnya melaksanakan penilaian pada proses pembelajaran. • Lembaga penyelenggara program kursus Bahasa Inggris seharusnya melakukan pengawasan proses pembelajaran. • Lembaga penyelenggara program kursus Bahasa Inggris seharusnya melakukan evaluasi pada akhir pendidikan. • Hasil pengawasan dan evaluasi proses pembelajaran sebaiknya dilaporkan kepada pihak-pihak yang terkait. 	Sesuai dengan Standar Nasional Pendidikan (SNP) Peraturan Pemerintah Republik Indonesia Nomor 19 tahun 2005 tentang standar proses evaluasi lembaga kursus dan pelatihan	
<i>Product</i>	Hasil belajar	<p><i>Tes</i></p> <ul style="list-style-type: none"> • Peserta didik memiliki kriteria nilai standar kelulusan yaitu 60 yang ditentukan dari pengetahuan, ketrampilan dan sikap (<i>listening, grammar, reading, writing, speaking</i>) 	Sesuai dengan standar penilaian yang ditentukan lembaga	Baik

3.8 Instrumen Penelitian

Kriteria Evaluasi

Untuk mengetahui kriteria evaluasi menggunakan perhitungan sebagai berikut:

1. Menjumlah skor semua subjek penelitian
2. Mencari nilai rata-rata (*Mean*) dan standar devisiasi (St.Dev)

Untuk mencari nilai rata-rata (Mean)

$$X = \frac{\sum X}{N}$$

Jadi, untuk mencari nilai rata-rata tinggal menjumlah semua skor, kemudian dibagi dengan banyaknya siswa yang memiliki skor itu. Sedangkan untuk mencari Standar Devisiasi:

$$St.Dev = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2}$$

Keterangan :

St.Dev = standar devisiasi

$\frac{\sum X^2}{N}$ = tiap skor dikuadratkan lalu dijumlahkan kemudian N dibagi N

$\left(\frac{\sum X}{N}\right)^2$ = semua skor dijumlahkan, dibagi N lalu dikuadratkan

(Arikunto, 2010:300)

3. Mengkonversi Skor Mentah Menjadi Nilai A, B, C, D dan EKelompok atas
Pengubahan skor mentah hasil tes menjadi nilai standar berskala lima (stanfive)
atau nilai huruf menggunakan patokan sebagai berikut:

$$M + 1,5 SD = A$$

$$M + 0,5 SD \text{ (sampai dengan) } M + 1,5 = B$$

$$M - 0,5 SD \text{ (sampai dengan) } M + 0,5 = C$$

$$M - 0,5 SD \text{ (sampai dengan) } M - 1,5 = D$$

$$M - 1,5 SD > E$$

Kemudian skor kohesivitas dikonversi menjadi beberapa tingkat kriteria yaitu: amat baik, baik, cukup, kurang, kurang sekali.

Tabel 3.3 Kriteria Evaluasi CIPP Pada Program Pembelajaran Bahasa Inggris Kelas *English for Children* di *English Smart* Bandar Jaya

Komponen	Sub Komponen	Skor	Kategori	Kriteria
<i>Context</i>	Kondisi Awal	23 ke atas 20-22 18-19 16-17 15 ke bawah	Amat Baik Baik Cukup Kurang Kurang Sekali	Amat Baik
<i>Input</i>	Sarana dan Prasarana	24 ke atas 23 22 21 20 ke bawah	Amat Baik Baik Cukup Kurang Kurang Sekali	Amat Baik
	Tenaga Pendidik	9 ke atas 8 7 6 5 ke bawah	Amat Baik Baik Cukup Kurang Kurang Sekali	Amat Baik
	Kurikulum	9 ke atas 8 7 6 5 ke bawah	Amat Baik Baik Cukup Kurang Kurang Sekali	Amat Baik
<i>Process</i>	Perencanaan Proses Pembelajaran	26 ke atas 25 21-24 20 19 ke bawah	Amat Baik Baik Cukup Kurang Kurang Sekali	Amat Baik
	Pelaksanaan Proses Pembelajaran	86 ke atas 83-85 80-82 77-79 76 ke bawah	Amat Baik Baik Cukup Kurang Kurang Sekali	Amat Baik
	Penilaian Proses Pembelajaran	25 ke atas 24 21-23 30 19 ke bawah	Amat Baik Baik Cukup Kurang Kurang Sekali	Amat Baik
<i>Product</i>	Hasil Belajar	84 ke atas 76-83 67-75 59-66 58 ke bawah	Amat Baik Baik Cukup Kurang Kurang Sekali	Baik

Sumber: Perhitungan hasil analisis komponen *context*, *input*, *process* dan *product* berdasarkan nilai Mean dan Standar Deviasi

3.9 Kisi-Kisi Instrumen Penelitian

Penyusunan instrumen berangkat dari kisi-kisi evaluasi yang aspek-aspek penilaiannya disesuaikan dengan ruang lingkup variabel yang akan diukur. Adapun kisi-kisi instrumen penelitian ini sebagai berikut:

Tabel 3.4 Kisi-Kisi Instrumen Penelitian

Dimensi	Aspek	Indikator	Teknik pengumpulan data	Subjek Penelitian	Jumlah Pertanyaan
context	Kondisi awal	1. Visi Misi 2. Pengelolaan 3. Kepemimpinan 4. Sistem informasi manajemen	Observasi, dokumentasi	Kepala lembaga, tentor	6
Input	Sarana dan prasarana	1. Ruang belajar 2. Ruang staff 3. Buku	Observasi, dokumentasi	Kepala lembaga, <i>supervisor</i> , tentor	12
	Tenaga Pendidik	Jenjang pendidikan			
	Kurikulum	Penggunaan kurikulum			
Process	Perencanaan	1. Kalender pendidikan 2. Silabus 3. <i>Lesson Plan</i> 4. Beban belajar 5. Bahan ajar	Observasi, dokumentasi	Tentor	10
	Pelaksanaan	1. Metode 2. Media	Observasi	Tentor	21
	Penilaian	1. <i>Pre-test</i> 2. <i>Post-test</i> 3. Laporan	Observasi	Tentor	6
Product	Hasil belajar	1. Pengetahuan 2. Ketrampilan 3. Sikap	Dokumentasi	Pembelajar	

3.10 Validitas dan Realibilitas Instrumen

Sebelum instrumen digunakan sebagai alat ukur evaluasi program pembelajaran Bahasa Inggris terlebih dahulu diuji coba validitasnya kepada responden di luar subjek uji coba, yaitu menggunakan para ahli yang berpendidikan minimal S2. Widoyoko (2012:141) menjelaskan bahwa instrumen dikatakan valid apabila instrumen tersebut dapat dengan tepat mengukur apa yang hendak diukur. Dengan kata lain validitas berkaitan dengan “ketepatan” dengan alat ukur. Dengan instrumen yang valid akan menghasilkan data yang valid pula.

Melalui penilaian terhadap kelayakan tampilan butir-butir pertanyaan, kemudian analisis yang lebih dalam dilakukan dengan maksud untuk menilai kelayakan isi butir pertanyaan sebagai jabaran dari indikator berperilaku atribut yang diukur. Penilaian ini bersifat kualitatif dan *judgemental* dan dilaksanakan oleh suatu panel *expert*, bukan oleh penulis butir. Inilah prosedur yang menghasilkan validitas logis (*logical validity*). Seberapa tinggi kesepakatan antara *expert* yang melakukan penilaian kelayakan suatu butir akan dapat diestimasi dan dikuantifikasikan, kemudian statistiknya dijadikan indikator validitas isi butir pertanyaan pada instrumen.

Penilaian ahli/pakar dilakukan dengan maksud untuk mengetahui validitas konstruk dari skala penilaian yang telah dikembangkan. Sasaran penilaian mencakup adanya kesesuaian penjabaran konstruk yang digunakan hingga menjadi butir instrumen.

Terhadap dua hal pokok yang dinilai oleh ahli/pakar, yaitu: a) kesesuaian indikator yang akan dikembangkan terhadap konsep atau konstruk yang digunakan, b) kesesuaian butir-butir instrumen yang akan dikembangkan terhadap indikator yang menjadi acuannya. Rancangan evaluasi yang telah dikonstruksi ini diajukan kepada 2 orang pakar/ahli.

Dalam hal ini pemberian skor pada jawaban setiap butir dengan menggunakan Skala Guttman, dimodifikasi sesuai dengan apa yang tercantum dalam penilaian.

Tabel 3.5 Skor Instrumen Evaluasi CIPP berdasarkan Skala Guttman

No.	Jawaban Item Instrumen Metainstrumen SBD_BK	Skor
1	YA (Sesuai/Terkait/Tepat/Jelas/Layak)	1
2	TIDAK (Tidak Sesuai/ Tidak Terkait/ Tidak Tepat/ Tidak Jelas/ Tidak Layak)	0

Tafsiran presentase digunakan untuk mengetahui banyaknya ahli/pakar yang memberikan respon. Tafsiran menurut Koentjaraningrat (dalam Ohira, 2013:15) adalah sebagai berikut :

Tabel 3.6 Tafsiran Persentase Penilaian

Rentang Persentase	Kategori
0	Tidak ada
1 – 25	Sebagian kecil
26 – 49	Hampir setengahnya
50	Setengahnya
51 - 75	Sebagian besar
76 – 99	Hampir seluruhnya
100	Seluruhnya

3.10.1 Hasil Uji Validitas Ahli

Pengujian validitas instrumen untuk pedoman evaluasi dilakukan oleh Gembong Sumadiyono, M.Pd (GS) dan Suharyanto, M.Pd (SH). Pada penelitian ini validitas isi pada umumnya melalui pertimbangan para ahli. Uji validitas isi tidak ada formula matematis untuk menghitung dan tidak ada cara untuk menunjukkan secara pasti. Tetapi untuk memberikan gambaran bagaimana suatu tes divalidasi dengan menggunakan validitas isi, pertimbangan ahli tersebut dilakukan dengan cara sebagai berikut : para ahli, pertama diminta untuk mengamati secara cermat semua butir dalam tes yang hendak divalidasi. Kemudian mereka diminta untuk mengoreksi semua butir-butir yang telah dibuat. Dan pada akhir perbaikan, mereka juga diminta untuk memberikan pertimbangan tentang bagaimana tes tersebut menggambarkan cakupan isi yang hendak diukur.

1. Hasil validitas ahli untuk instrumen evaluasi *Context*

Berdasarkan data telaah pakar, diketahui bahwa:

Tabel 3.7 Hasil Validitas Telaah Ahli/Pakar Pada Instrumen *Context*

No.	Ahli/Pakar	Persentase
1.	GS	87,50%
2.	SH	87,50%
Rata-rata		87,50%

Berdasarkan penilaian dari pakar sebagaimana dijelaskan pada data hasil uji coba teoretik, terlihat bahwa persentase jawaban YA berada pada rentang 76 – 99%. Rentang persentase ini dalam kategori hampir seluruh butir pernyataan dapat

digunakan untuk mengungkapkan kualitas perencanaan pembelajaran Bahasa Inggris. Skala penilaian perencanaan pembelajaran Bahasa Inggris dianggap telah sesuai dengan teori dan mampu mengukur apa yang hendak diukur. Berdasarkan hasil uji validitas yang dilakukan oleh para ahli, maka pada instrumen skala penilaian kategori pelaksanaan pembelajaran Bahasa Inggris perlu dilakukan perbaikan, diantaranya:

- a. Pernyataan no. 1 dan 4 disederhanakan
- b. Pernyataan no. 2 pernyataannya direvisi agar mudah di pahami

2. Hasil validitas ahli untuk instrumen evaluasi input

Berdasarkan data telaah pakar, diketahui bahwa:

Tabel 3.8 Hasil Validitas Telaah Ahli/Pakar Pada Instrumen Input.

No.	Ahli/Pakar	Persentase
1.	GS	93,70%
2.	SH	87,50%
Rata-rata		90,60%

Berdasarkan penilaian dari pakar sebagaimana dijelaskan pada data hasil uji coba teoretik, terlihat bahwa persentase jawaban YA berada pada rentang 76 – 99%. Rentang persentase ini dalam kategori hampir seluruh butir pernyataan dapat digunakan untuk mengungkapkan kategori pelaksanaan pembelajaran Bahasa Inggris. Skala penilaian pelaksanaan pembelajaran Bahasa Inggris dianggap telah sesuai dengan teori dan mampu mengukur apa yang hendak diukur. Dengan demikian skala penilaian pelaksanaan pembelajaran Bahasa Inggris perlu dilakukan revisi secara keseluruhan. Berdasarkan hasil uji validitas yang dilakukan oleh para ahli, maka pada

instrumen skala penilaian kategori pelaksanaan pembelajaran Bahasa Inggris perlu dilakukan perbaikan, diantaranya:

- a. Pernyataan no. 1 disederhanakan agar mudah dipahami
- b. Pernyataan no. 2 direvisi
- c. Pernyataan no.2 dalam indikator di ganti karena tidak sesuai

3. Hasil validitas ahli untuk instrumen evaluasi proses

Berdasarkan data telaah pakar, diketahui bahwa:

Tabel 3.9 Hasil Validitas Telaah Ahli/Pakar Pada Instrumen Proses

No.	Ahli/Pakar	Persentase
1.	GS	93,75%
2.	SH	93,75%
Rata-rata		93,75%

Berdasarkan penilaian dari pakar sebagaimana dijelaskan pada data hasil uji coba, terlihat bahwa persentase jawaban YA berada pada rentang 76-99%. Rentang persentase ini dalam kategori hampir seluruh butir pernyataan dapat digunakan untuk mengungkapkan kategori evaluasi pembelajaran Bahasa Inggris. Skala penilaian evaluasi pembelajaran Bahasa Inggris dianggap telah sesuai dengan teori dan mampu mengukur apa yang hendak diukur. Tetapi perlu dilakukan perevisian pada beberapa butir. Berdasarkan hasil uji validitas yang dilakukan oleh para ahli, maka pada instrumen skala penilaian evaluasi pembelajaran Bahasa Inggris perlu dilakukan perbaikan, yaitu:

- a. Pernyataan no. 2 direvisi
- b. Pernyataan no. 2 dan 6 pada sub indikator pelaksanaan pembelajaran direvisi

- c. Pernyataan indikator no. 3 direvisi pada kata “remedi”

3.10.2 Hasil Uji Realibilitas Ahli

Untuk menguji realibilitas instrumen pembelajaran Bahasa Inggris tahap ujicoba teoretik dari para ahli/pakar digunakan *inter-rater reliability*, yaitu reliabilitas yang dilihat dari tingkat kesepakatan (*aggreement*) antara *rater* (penilai). *Inter-rater reliability* (IRR) akan memberikan gambaran (berupa skor) tentang sejauhmana tingkat konsensus atau kesepakatan yang diberikan ahli/pakar. Koefisien IRR yang digunakan adalah koefisien kesepakatan *Cohen Kappa* (K) dengan formula sebagai berikut Bhisma Murti dalam Ohira, (2013 : 18);

$$K = \frac{Po - Pe}{1 - Pe}$$

Keterangan:

K = Koefisien *Cohen Kappa*

Po = Proporsi Kesepakatan teramati

Pe = Proporsi kesepakatan harapan

1 = Konstanta

Hasil yang diperoleh dari penilaian ahli terhadap instrumen evaluasi pembelajaran Bahasa Inggris dianalisis secara kuantitatif dengan bantuan software SPSS (*Statistical Program for Social Science*) versi 17.0. Interpretasi kesepakatan *Kappa* yang dipakai adalah 0.61-0.80 (baik) menurut tabel interpretasi *Kappa* oleh Altman dalam Ohira (2013:18) sebagai berikut:

Tabel 3.10 Kekuatan Koefisien Kappa

Nilai Kappa	Kekuatan Kesepakatan
$\leq 0,20$	Buruk
0,20 – 0,40	Kurang dari sedang
0,41 – 0,60	Sedang
0,61 – 0,80	Baik
0,81 – 1,00	Sangat Baik

Berdasarkan data hasil metaskala penilaian perencanaan, reliabilitas skala penilaian perencanaan dianalisis dengan menggunakan analisis *inter-rater reliability* (IRR) koefisien *Cohens's Kappa* terhadap kesepakatan (*agreement*) 2 orang ahli/pakar. Setelah dilakukan analisis maka diketahui koefisien Kappa sebagaimana dalam tabel berikut:

Tabel 3.11 *Output* Reliabilitas Hasil Analisis Uji Coba Instrumen Evaluasi *Context*
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Ahli1_context * Ahli2_context	8	100.0%	0	.0%	8	100.0%

Ahli1_context * Ahli2_context Crosstabulation

Count

		Ahli2_context		Total
		1.00	2.00	
Ahli1_context	1.00	0	3	3
	2.00	2	3	5
Total		2	6	8

Symmetric measure

		Value	Asymp. Std. Err ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement	Kappa	.429	.202	-1.265	.206
N of Valid Cases		8			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Berdasarkan analisis SPSS seri 17.0, nilai value instrumen evaluasi *context* diatas adalah 0,429. Berdasarkan tabel kriteria reliabilitas nilai 0,41 – 0,60 berarti data tersebut dikategorikan memiliki nilai reliabilitas yang sedang sesuai dengan kekuatan koefisien kappa. Maka data instrumen evaluasi *context* dikatakan reliable dan dapat digunakan untuk penelitian yang sebenarnya di lapangan. *Asymp Std. Error* menunjukkan kesalahan pengukuran terstandar, semakin kecil besarnya koefisien ini, semakin reliabel hasil pengukuran yang dihasilkan.

Tabel 3.12 *Output* Reliabilitas Hasil Analisis Uji Coba Instrumen Evaluasi *Input*
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Ahli1_input * Ahli2_input	10	100.0%	0	.0%	10	100.0%

Ahli1_context * Ahli2_context Crosstabulation

Count

		Ahli2_input		Total
		1.00	2.00	
Ahli1_input	1.00	0	3	3
	2.00	3	4	7
Total		3	7	10

Symmetric Measures

		Value	Asymp. Std. Err ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement	Kappa	.701	.202	-1.265	.206
N of Valid Cases		8			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Berdasarkan analisis SPSS seri 17.0, nilai value instrumen devaluasi *input* di atas adalah 0,70. Berdasarkan tabel kriteria reliabilitas nilai 0,61 – 0,80 berarti data tersebut dikategorikan memiliki nilai reliabilitas yang baik sesuai dengan kekuatan koefisien kappa. Maka data instrumen evaluasi *input* dikatakan reliabel dan dapat digunakan untuk penelitian yang sebenarnya di lapangan. *Asymp Std. Error* menunjukkan kesalahan pengukuran terstandar, semakin kecil besarnya koefisien ini, semakin reliabel hasil pengukuran yang dihasilkan.

Tabel 3.13 *Output* Reliabilitas Hasil Analisis Uji Coba Instrumen Evaluasi *Process*
Case Processing Summary

		Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
rocess * Ahli2_rocess	25	100.0%	0	.0%	25	100.0%

Ahli1_process * Ahli2_process Crosstabulation

Count

		Ahli2_process		Total
		1.00	2.00	
Ahli1_process	1.00	0	2	2
	2.00	3	20	23
Total		3	22	25

Symmetric Measures

		Value	Asymp. Std. Err ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement	Kappa	.610	.052	-.544	.586
N of Valid Cases		25			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Berdasarkan analisis SPSS seri 17.0, nilai value instrumen evaluasi *process* di atas adalah 0.610. berdasarkan tabel kriteria reliabilitas nilai 0.61-0.80 berarti data tersebut dikategorikan memiliki nilai reliabilitas yang baik sesuai kekuatan koefisien kappa. Maka data instrumen evaluasi *process* dikatakan reliable dan dapat digunakan untuk penelitian yang sebenarnya di lapangan. *Asymp Std. Error* menunjukkan kesalahan pengukuran terstandar, semakin kecil besarnya koefisien ini, semakin reliabel hasil pengukuran yang dihasilkan.