

**VALIDITAS DAN RELIABILITAS INSTRUMEN PENILAIAN  
PROGRAM BAHASA INGGRIS GRUP 5 FAKULTAS EKONOMI  
BISNIS ISLAM PUSAT PENGEMBANGAN BAHASA  
IAIN PADANGSIDIMPUAN  
T.P. 2018/2019**

<sup>1</sup>Zulpan, <sup>2</sup>Rosmiyasari Ritonga

<sup>1</sup>Universitas Islam Labuhan Batu

<sup>2</sup>Universitas Islam Negeri Sultan Syarif Kasim

<sup>1</sup>email: zulpan200990@gmail.com

<sup>2</sup>email: rosmia2sari@gmail.com

<sup>1</sup>Jl. H. M. Yunus No. 09 Padang Bulan, Kabupaten *Labuhan Batu*, Sumatera Utara.

<sup>2</sup>Jl. H.R Soebrantas No 155 KM.15 Simpang Baru Panam Pekanbaru

***Abstract***

*This research is aimed to describe validity and reliability of instrument on English assessment at Language Centre of IAIN Padangsidimpuan Program. The descriptive quantitative is kind of research method that used on this research. Subject of this research is the group 5 of Economy of Islamic Business Faculty at IAIN Padangsidimpuan that consisted with 24 students. Data collected from result of reading, writing, listening, speaking, and structure of the students. Method of analysis data is taken by using validity and reliability tests through software SPSS version 16. The result showed that all of indicators on mid-semester test were valid and the instrument is reliable meanwhile in the final-semester test showed reading, writing, listening, speaking were not valid and the structure one was valid and the instruments were not reliable.*

**Kata-kata Kunci: Instrumen, Penilaian, Reliabelitas, Validitas.**

***Abstrak***

Penelitian ini bertujuan untuk mendeskripsikan validitas dan reliabilitas instrumen penilaian bahasa Inggris pada program Lembaga Pusat Pengembangan Bahasa IAIN Padangsidimpuan. Metode penelitian menggunakan metode penelitian kuantitatif deskriptif. Subjek penelitian ini adalah grup 5 Fakultas Ekonomi Bisnis Islam IAIN Padangsidimpuan yang terdiri dari 24 mahasiswa. Data yang dikumpulkan adalah hasil tes *Reading, Writing, Listennig, Speaking and Structure* mahasiswa dilakukan. Metode analisis data menggunakan uji validitas dan reliabilitas berbantuan software SPSS versi 16. Hasil penelitian menunjukkan pada MID semester semua indikator dikatakan valid dan instrumen reliabel sementara pada FINAL *Reading, Writing, Listennig, Speaking* dikatakan tidak valid dan hanya *Structure* yang valid dan instrumen tidak reliabel.

**Keywords: Assessment, Instrument, Realibility, Validity.**

## 1. PENDAHULUAN

IAIN Padangsidimpuan memiliki program unggulan yaitu program Ma'Had Al-Jamiah dengan penekanan wajib berbahasa Arab dan Inggris di arena kampus dan asrama. Wajib berbahasa Inggris merupakan jembatan untuk dapat berkomunikasi di penjuru dunia walaupun berbeda negara. Dengan adanya program tersebut diharapkan seluruh masyarakat baik dosen dan mahasiswa di ruang lingkup IAIN Padangsidimpuan secara berkelanjutan dapat menerapkan dan menggunakan bahasa Inggris guna untuk memiliki skill berbahasa Inggris yang baik.

Pentingnya belajar bahasa Inggris harus ditanamkan dari pribadi kita sendiri, hal tersebut dapat menjadikan kita berkembang secara nasional dan internasional untuk dapat memperoleh pengetahuan dan ilmu yang luas. Bahasa Inggris adalah bahasa pengantar dimana hasil karya, terbaharuan ilmu di seluruh dunia telah di tuangkan dalam bahasa Inggris baik tertulis dan lisan. Maka dengan itu kita harus menyadari perlunya belajar yang maksimal untuk memperoleh hasil yang maksimal khususnya dalam kemampuan berbahasa Inggris.

Hasil belajar merupakan tolak ukur kita untuk mengetahui sejauh mana capaian yang telah kita miliki dalam belajar. Hasil belajar merupakan kemampuan yang diperoleh individu setelah proses belajar berlangsung, yang dapat memberikan perubahan tingkah laku baik pengetahuan, pemahaman, sikap dan keterampilan siswa sehingga menjadi lebih baik dari sebelumnya. Sebagaimana yang dikemukakan (Sanjaya, 2010) hasil belajar adalah suatu penilaian akhir dari proses dan pengenalan yang telah dilakukan berulang-ulang serta akan tersimpan dalam jangka waktu lama atau bahkan tidak akan hilang selama-lamanya karena hasil belajar turut serta dalam membentuk pribadi individu yang selalu ingin mencapai hasil yang lebih baik lagi sehingga akan merubah cara berpikir serta menghasilkan perilaku kerja yang lebih baik (Sjukur, 2012).

Proses evaluasi suatu tindakan yang digunakan untuk memperoleh hasil yang didapatkan pada saat proses belajar. Penilaian, pengukuran, tes merupakan bagian penting dalam proses evaluasi hasil belajar. Penilaian menurut Permendikbud Tahun 2016 adalah proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik. Penilaian adalah sebagai proses sistematis untuk mengumpulkan, menganalisis, menginterpretasikan informasi, dan menentukan tingkat keberhasilan siswa terhadap tujuan pembelajaran (Rosilawati, 2017). Daryanto & Amirono (2016) menjelaskan penilaian (*assessment*) adalah penerapan berbagai cara penggunaan beragam alat penilaian untuk memperoleh tentang sejauh mana hasil belajar peserta didik atau ketercapaian kompetensi (rangkaiian kemampuan) peserta didik. Penilaian adalah mengambil suatu keputusan, terhadap sesuatu dengan mengacu kepada ukuran tertentu seperti baik dan buruk, pandai atau bodoh, tinggi dan rendahnya dan sebagainya (Supardi, 2015). Proses penilaian meliputi pengumpulan bukti- bukti tentang pencapaian belajar peserta didik. Bukti ini tidak selalu diperoleh melalui tes saja, tetapi juga bisa dikumpulkan melalui pengamatan atau laporan diri (Mardapi, 2016).

Tujuan penilaian sebagai berikut: (1) penilaian hasil belajar oleh pendidik bertujuan untuk memantau dan mengevaluasi proses, kemajuan belajar, dan perbaikan hasil belajar peserta didik secara berkesinambungan. (2) penilaian hasil belajar oleh satuan pendidikan bertujuan untuk menilai pencapaian Standar Kompetensi Lulusan untuk semua mata pelajaran. (3) penilaian hasil belajar oleh pemerintah bertujuan untuk menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu. Daryanto & Amirono (2016) menyatakan tujuan penilaian sebagai berikut: (1) Mendeskripsikan kecakapan belajar para peserta didik sehingga dapat diketahui kelebihan dan kekurangannya dalam berbagai bidang studi atau mata pelajaran yang ditempuhnya. (2) Mengetahui keberhasilan proses pendidikan dan pengajaran di sekolah, yakni seberapa jauh keefektifannya dalam mengubah tingkah laku para peserta didik ke arah tujuan pendidikan yang diharapkan. (3) Menentukan tindak lanjut hasil penilaian, yakni melakukan perbaikan dan penyempurnaan dalam hal program pendidikan dan pengajaran serta strategi pelaksanaannya. (4) Memberikan pertanggungjawaban (*accountability*) dari pihak sekolah kepada pihak-pihak yang berkepentingan.

Instrumen adalah alat yang digunakan untuk pengumpulan data, alat yang dilakukan yaitu tes dan nontes. Dalam hal ini Lembaga Pusat Pengembangan Bahasa IAIN Padangsidimpuan telah memiliki tes maupun nontes pada pengembangan bahasa Inggris yaitu instrumen *Reading, Writing, Listennig, Speaking and Structure*. Hal ini peneliti memandang penting dari dasar teoritis menjadi komponen-komponen penilaian, dimana hal ini berarti peneliti harus melihat validitas, reliabilitas guna menghasilkan instrumen yang baik. Menganalisis instrumen *reading, writing, Listennig, Speaking and Structure* bertujuan untuk mengetahui apakah alat ukur yang digunakan atau yang akan digunakan sudah memenuhi syarat-syarat sebagai alat ukur yang baik, tepat mengukur sesuatu sesuai dengan tujuan yang telah dirumuskan. Sebuah instrumen dikatakan baik jika memenuhi syarat validitas, reliabilitas baik.

Validitas berasal dari kata *validity* yang mempunyai arti sejauh mana akurasi suatu tes atau skala dalam menjalankan fungsi pengukurannya (Azwar, 2012). Ada lima sumber bukti validitas yang penting yaitu, bukti berdasarkan isi tes, bukti berdasarkan proses respons, bukti berdasarkan struktural internal, bukti berdasarkan hubungan dengan variabel lain, dan bukti berdasarkan hubungan dengan variabel lain, dan bukti berdasarkan konsekuensi pengujian. Validitas isi adalah merupakan yang diestimasi lewat pengujian terhadap kelayakan atau relevansi isi tes melalui analisis rasional oleh panel berkompeten atau melalui *expert judgment* (Azwar, 2012). Sedangkan validitas konstruk adalah validitas yang menunjukkan sejauhmana hasil tes mampu mengungkap suatu *trait* atau suatu konstruk teoritik yang hendak di ukur (Allen & Yen, 1979). Validitas konstruk merupakan jenis validitas kuantitatif yang didasarkan data empirik lapangan dengan angka statistik yang memberikan informasi tentang validitas data (Periantalo, 2015). Reliabilitas adalah sejauhmana hasil suatu proses pengukuran dapat di percaya. Reliabilitas mempunyai berbagai nama lain seperti konsistensi, keterandalan, keterpercayaan, kestabilan,

keajegan dan sebagainya. Reliabilitas atau keandalan merupakan koefisien yang menunjukkan tingkat keajegan atau konsistensi hasil pengukuran suatu tes (Mardapi, 2016). Reliabilitas mengacu kepada konsistensi atau keakuratan hasil ukur. Konsistensi hasil pengukuran dengan menggunakan alat ukur yang sama untuk orang yang berbeda atau pada waktu yang berbeda tetapi kondisi yang sama. Sebuah instrumen dikatakan baik jika memenuhi syarat validitas, reliabilitas yang baik. Penelitian ini dilakukan untuk melihat validitas dan reliabilitas instrumen *Reading, Writing, Listennig, Speaking and Structure* pada Lembaga Pusat Pengembangan Bahasa IAIN Padangsidempuan khususnya grup 5 Fakultas Ekonomi Bisnis Islam IAIN Padangsidempuan.

## 2. METODE PENELITIAN

Metode penelitian yang digunakan adalah metode penelitian kuantitatif deskriptif, dimana penelitian mendeskripsikan hasil validitas butir dan reliabilitas instrumen yang digunakan di Lembaga Bahasa IAIN Padangsidempuan. Subjek penelitian ini dilakukan di grup 5 Fakultas Ekonomi Bisnis Islam IAIN Padangsidempuan terdiri dari 24 mahasiswa. Pengumpulan data yang digunakan dalam penelitian ini adalah hasil tes *Reading, Writing, Listennig, Speaking and Structure* mahasiswa dilakukan. Metode analisis data menggunakan uji validitas terhadap indikator kemampuan berbahasa yaitu; *Reading, Writing, Listennig, Speaking and Structure* mahasiswa sementara uji reliabilitas instrumen kemampuan berbahasa Inggris di Lembaga Pusat Pengembangan Bahasa IAIN Padangsidempuan secara holistik. Hal tersebut digunakan berbantuan software SPSS versi 16.

## 3. HASIL PENELITIAN

Validitas digunakan untuk mengetahui sejauhmana hasil pengukuran dari indikator-indikator dalam instrumen yang digunakan Lembaga Bahasa IAIN Padangsidempuan. Hasil uji validitas instrumen dapat ditentukan ketentuan jika nilai hitung  $> r$  tabel maka instrumen dapat dikatakan Valid dan jika nilai hitung  $< r$  tabel maka instrumen dapat dikatakan tidak Valid. Validitas instrumen dilihat berdasarkan hasil pengukuran setiap indikator dapat dilihat sebagai berikut:

**Tabel 1: Hasil Uji Validitas Instrumen MID SEMESTER**

Indikator	r hitung		r tabel Sig 5%	Kategori
Reading	0.541	>	0.404	Valid
Writing	0.588	>	0.404	Valid
Listening	0.601	>	0.404	Valid
Speaking	0.662	>	0.404	Valid
Structure	0.885	>	0.404	Valid

**Tabel 2: Hasil Uji Validitas Instrumen Final SEMESTER**

Indikator	r hitung		r tabel Sig 5%	Kategori
Reading	0.297	<	0.404	Tidak Valid
Writing	0.359	<	0.404	Tidak Valid
Listening	-0.205	<	0.404	Tidak Valid
Speaking	0.344	<	0.404	Tidak Valid
Structure	0.656	>	0.404	Valid

Dari hasil analisis diatas tersebut pada hasil MID semester semua indikator penilaian *Reading, Writing, Listennig, Speaking and Structure* adalah valid sementara pada hasil FINAL semester indikator penilaian *Reading, Writing, Listennig, Speaking* dikatakan tidak valid dan hanya indikator penilaian *Structure* yang valid.

Hasil uji reliabilitas intrumen dapat ditentukan sebagai berikut (Ghozali 2002) jika nilai *cronbach's alpha* ( $\alpha$ ) > 0,6, maka instrumen dapat dikatakan reliabel dan jika nilai *cronbach's alpha* ( $\alpha$ ) < 0,6, maka instrumen r dapat dikatakan tidak reliabel, dilihat sebagai berikut:

**Tabel 3: Hasil Uji Reliabilitas Instrumen MID SEMESTER**

Hasil hitung		$\alpha$	Kategori
0.611	>	0.6	Reliabel

**Tabel 4: Hasil Uji Reliabilitas Instrumen FINAL SEMESTER**

Hasil hitung		$\alpha$	Kategori
-0.978	<	0.6	Tidak Reliabel

Dari hasil analisis diatas tersebut pada hasil MID semester instrumen penilaian kemampuan berbahasa Inggris dikatakan reliabel sementara pada hasil FINAL semester instrumen penilaian kemampuan berbahasa Inggris dikatakan tidak reliabel.

#### 4. KESIMPULAN

Berdasarkan hasil analisis validitas instrumen terhadap indikator penilaian serta reliabilitas instrumen yang dilakukan oleh Lembaga Pusat Pengembangan Bahasa IAIN Padangsidimpuan pada kemampuan berbahasa Inggris di MID semester dikatakan valid dan reliabel sementara pada FINAL *Reading, Writing, Listennig, Speaking* dikatakan tidak valid dan hanya *Structure* yang valid dan instrumen tidak reliabel. Hal tersebut bisa menjadi evaluasi sekaligus refleksi untuk menyelidiki dimana letak kesalahan penilaian pada FINAL semester, secara teori dapat dilihat dari faktor-faktor internal mahasiswa atau dosen, konstruksi soal, konstruksi indikator soal dan sebagainya.

**DAFTAR PUSTAKA**

- Allen, M.J. & Yen W.M. (1979). *Introduction to Mesurement Theory*. Monterey:Book/Cole.
- Azwar, S. (2012). *Reliabilitas dan Validitas edisi 4*. Yogyakarta: Pustaka Pelajar.
- Azwar, S. (2016). *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Belajar.
- Daryanto & Amirono. (2016). *Evaluasi dan Penilaian Pembelajaran Kurikulum 2013*. Yogyakarta: Gava Media.
- Mardapi, D. (2016). *Pengukuran Penilaian & Evaluasi Pendidikan*. Yogyakarta: Parama Publishing. <https://doi.org/978-602-6243-20-1>
- Nasehudin, T. S., & Gozali, N. (2012). *Metode Penelitian Kuantitatif*. Bandung: Pustaka Setia.
- Periantalo, J. (2015). *Validitas Alat ukur Psikologi: Aplikasi Praktis*. Yogyakarta: Pustaka Belajar.
- Permendikbud. Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 23 Tahun 2016 Tentang Standar Penilaian Pendidikan, Pub. L. No. 897 (2016). Indonesia.
- Retnawati, H. (2016). *Validitas Reliabilitas & karakteristik Butir*. Yogyakarta: Parama Publishing.
- Rosilawati, A. (2017). *Pengembangan Instrumen Penilaian*. Semarang: Unnes Press.
- Sanjaya, Wina (2010). *Kurikulum dan pembelajaran: Teori dan praktik pengembangan kurikulum tingkat satuan pendidikan (KTSP)*. Jakarta: Kencana Prenada Media Group.
- Sjukur, Sulihin B. (2012) Pengaruh Blended Learning Terhadap Motivasi Belajar Dan Hasil Belajar Siswa Tingkat SMK. *Jurnal Pendidikan Vokasi*, Vol 2, Nomor 3.
- Sugiyono. (2016). *Metode Penelitian Kuantitatif, Kualitatif, dan Kombinasi (Mixed Method)*. Bandung: PT. Alfabeta.
- Supardi. (2015). *Penilaian Autentik Pembelajaran Afektif, Kognitif, Psikomotorik (Konsep dan Aplikasi)*. Jakarta: Rajawali pers.

## Lampiran

## Hasil Output Validitas dan Reliabilitas MID Semester

		Correlations					SKOR_TOTAL
		READING	WRITING	LISTENING	SPEAKING	STRUCTURE	
READING	Pearson Correlation	1	-.136	.080	.224	.498*	.541**
	Sig. (2-tailed)		.526	.712	.294	.013	.006
	N	24	24	24	24	24	24
WRITING	Pearson Correlation	-.136	1	.641**	.528**	.295	.588**
	Sig. (2-tailed)	.526		.001	.008	.162	.002
	N	24	24	24	24	24	24
LISTENING	Pearson Correlation	.080	.641**	1	.479*	.329	.601**
	Sig. (2-tailed)	.712	.001		.018	.117	.002
	N	24	24	24	24	24	24
SPEAKING	Pearson Correlation	.224	.528**	.479*	1	.364	.662**
	Sig. (2-tailed)	.294	.008	.018		.080	.000
	N	24	24	24	24	24	24
STRUCTURE	Pearson Correlation	.498*	.295	.329	.364	1	.885**
	Sig. (2-tailed)	.013	.162	.117	.080		.000
	N	24	24	24	24	24	24
SKOR_TOTAL	Pearson Correlation	.541**	.588**	.601**	.662**	.885**	1
	Sig. (2-tailed)	.006	.002	.002	.000	.000	
	N	24	24	24	24	24	24

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

## Reliability Statistics

Cronbach's Alpha	N of Items
.611	5

## Hasil Output Validitas dan Reliabilitas FINAL Semester

		Correlations					SKOR_TOTAL
		READING	WRITING	LISTENING	SPEAKING	STRUCTURE	
READING	Pearson Correlation	1	-.177	.124	.117	-.384	.297
	Sig. (2-tailed)		.407	.564	.586	.064	.159
	N	24	24	24	24	24	24
WRITING	Pearson Correlation	-.177	1	-.281	.458*	.167	.359
	Sig. (2-tailed)	.407		.184	.025	.435	.085
	N	24	24	24	24	24	24
LISTENING	Pearson Correlation	.124	-.281	1	-.297	-.455*	-.205
	Sig. (2-tailed)	.564	.184		.158	.026	.337
	N	24	24	24	24	24	24
SPEAKING	Pearson Correlation	.117	.458*	-.297	1	-.039	.344
	Sig. (2-tailed)	.586	.025	.158		.857	.100
	N	24	24	24	24	24	24
STRUCTURE	Pearson Correlation	-.384	.167	-.455*	-.039	1	.656**
	Sig. (2-tailed)	.064	.435	.026	.857		.000
	N	24	24	24	24	24	24
SKOR_TOTAL	Pearson Correlation	.297	.359	-.205	.344	.656**	1
	Sig. (2-tailed)	.159	.085	.337	.100	.000	
	N	24	24	24	24	24	24

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

## Reliability Statistics

Cronbach's Alpha <sup>a</sup>	N of Items
-.978	5

a. The value is negative due to a negative average covariance among items. This violates reliability model assumptions. You may want to check item codings.