

ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI RELIABILITAS TES TATA BAHASA INGGRIS DI UNIVERSITAS DI INDONESIA

Tri Jampi Setiyorini¹, Zyah Rochmad Jaelani², Abdul Ngafif³

Universitas Sarjanawiyata Tamansiswa Yogyakarta¹

Universitas Muhammadiyah Purworejo²

trijampi95@gmail.com¹, zyahjaelani@gmail.com², abdulngafif@gmail.com³

ABSTRAK

Penelitian ini bertujuan untuk menganalisis reliabilitas dan faktor-faktor yang mempengaruhi reliabilitas tes tata bahasa Inggris di sebuah universitas di Indonesia. Penelitian ini kualitatif, dan menggunakan studi kasus deskriptif. Instrumennya adalah tes, angket tertutup, dan wawancara terbuka. Berdasarkan hasil penelitian, reliabilitas tes adalah 0,754, sehingga tes termasuk reliabel. Dilihat dari hasil angket, faktor-faktor yang mempengaruhi reliabilitas tes adalah reliabilitas terkait mahasiswa, reliabilitas penilai, reliabilitas pelaksanaan tes, dan reliabilitas tes. Dilihat dari hasil wawancara, faktor-faktor yang mempengaruhi reliabilitas tes adalah penguasaan materi pembelajaran, persiapan belajar sebelum tes, kondisi fisik siswa, kejelasan petunjuk tes, kejelasan soal dalam tes, suasana dan alokasi waktu untuk melakukan tes.

Kata kunci: tata Bahasa Inggris, reliabilitas, tes

ABSTRACT

This study aims to analyse the reliability and the factors affecting the reliability of an English grammar test in an Indonesian university. This research was qualitative, and it used a descriptive case study. The instruments were tests, close-ended questionnaires, and open-ended interviews. Based on the research results, the reliability of the test is 0.754, so the test is reliable. Viewed from the questionnaires' results, the factors affecting the reliability of the test are student-related reliability, rater reliability, test administration reliability, and test reliability. Viewed from the interviews' results, the factors influencing the test's reliability are the mastery of learning material, the preparation of learning before getting the test, students' physical condition, the clarity of instructions of the test, clarity of questions in the test, the atmosphere and time allocation for doing the test.

Keywords: English grammar, reliability, test

PENDAHULUAN

Alat penilaian (testing) adalah cara untuk menentukan apakah mahasiswa dapat mencapai tujuan pembelajaran. Penilaian atau tes dapat menentukan keberhasilan program pembelajaran (Rixon dalam Apsari & Haryudin, 2017). Selain itu juga dapat

menginformasikan kepada mahasiswa tentang penguasaan materi pembelajarannya (Wenno et al., 2021). Penilaian adalah proses mengumpulkan informasi tentang prestasi mahasiswa untuk membuat keputusan yang baik (Ulfah et al., 2020). Alasan untuk menguji mahasiswa adalah untuk memberikan nilai kepada

mahasiswa, menentukan kompetensi mahasiswa untuk perbaikan yang sesuai, dan mengidentifikasi pengajaran yang tidak efektif. Dengan demikian, penilaian dapat membantu pendidik meningkatkan aktivitas pembelajaran (Reiser & Dick dalam Apsari & Haryudin, 2017).

Penilaian memberikan informasi tentang keputusan instruksional, mendeteksi kelebihan dan kelemahan mahasiswa dalam pembelajaran di kelas, dan memberi mereka umpan balik. Penilaian juga memberikan umpan balik langsung untuk meningkatkan praktik mengajar dosen. Dosen harus menggunakan tes untuk menentukan nilai mereka (Tosuncuoglu, 2018).

Tes memiliki peran penting dalam pendidikan di tingkat yang lebih tinggi, terutama di universitas. Tes memiliki peran penting dalam pendidikan untuk proses pengukuran dan evaluasi. Tes merupakan salah satu unsur penting dalam proses pembelajaran. Tes merupakan jenis evaluasi yang merupakan instrumen pengukuran proses pembelajaran. Ini merupakan instrumen untuk menilai keterampilan atau pengetahuan mahasiswa untuk mengambil keputusan pendidikan (Sugianto, 2017). Tes juga mengukur kemampuan bahasa mahasiswa

(Hughes dalam Setiabudi et al., 2019). Fungsi tes adalah untuk menilai kemampuan, pengetahuan, dan kinerja seseorang (Brown dalam Furwana, 2019). Tes dapat mengevaluasi keterampilan atau pengetahuan individu dalam standar tertentu. Dalam praktik pendidikan, tes menentukan kemampuan peserta didik untuk menyelesaikan tugas-tugas tertentu (Adom et al., 2020). Tes dapat dirancang dengan skor hasil yang diberikan kepada peserta didik untuk mendeteksi kemampuannya (Azmi, 2020).

Tes harus memenuhi ciri-ciri tes yang baik yaitu harus valid dan reliabel (Furwana, 2019). Tes yang baik adalah tes yang valid, reliabel, objektif, praktis, dan ekonomis (Djiwandono dalam Sugianto, 2017; Gyll & Ragland dalam Wenno et al., 2021). Dengan memberikan tes yang baik, mahasiswa berpeluang untuk mendapatkan kualitas belajar yang baik, dan hasilnya dapat meningkatkan proses belajar mengajar dan menentukan nilainya (Furwana, 2019).

Reliabilitas adalah konsistensi pengukuran dan stabilitas skor (Harris dalam Sultana, 2015). Artinya pengukuran yang sama menghasilkan hasil yang sama (Moser dan Kalton dalam Taherdoost, 2016). Jika dosen melakukan tes yang sama pada peserta yang sama pada

kesempatan yang berbeda, dan hasilnya serupa, maka tes itu reliable (Arifin, 2018; Heaton dalam Sultana, 2015). Prinsip reliabilitas adalah konsistensi skor, instruksi yang jelas, dan pertanyaan yang jelas (Tosuncuoglu, 2018). Interval antara pemberian dua tes tidak boleh terlalu lama atau terlalu pendek untuk meningkatkan reliabilitas (Hughes dalam Öz & Özturan, 2018). Reliabilitas adalah kualitas tes yang penting. Ketika konsistensi tes tercapai, validitas tes tercapai (Linn & Gronlund dalam Rosaroso, 2015). Untuk menghitung reliabilitas, diperlukan perhitungan statistik (James, 2013). Zimmerman & Zumbo (2015) mendefinisikan rasio reliabilitas varian nilai benar dan varian nilai teramati, di mana varian nilai teramati merupakan penjumlahan komponen benar dan salah. Hinton et al. dalam Taherdoost (2016) mengklasifikasikan reliabilitas menjadi reliabilitas sangat tinggi (0,90 ke atas), reliabilitas tinggi (0,70-0,90), reliabilitas sedang (0,50-0,70) dan reliabilitas rendah (0,50 ke bawah).

Ary et al. (2010) menyatakan bahwa ada tiga koefisien reliabilitas.

1. Reliabilitas test-retest

Reliabilitas test-retest adalah konsistensi hasil dari sampel yang sama pada waktu

yang berbeda (Ary et al., 2010). Ini adalah ukuran konsistensi dari sampel yang sama. Instrumen yang reliabel diharapkan dapat menghasilkan data yang serupa (James, 2013).

2. Bentuk alternatif atau bentuk reliabilitas paralel

Bentuk alternative atau reliabilitas paralel mengukur reliabilitas menggunakan dua bentuk instrumen. Memiliki domain yang sama, jumlah butir soal yang sama, spesifikasi tes yang sama, tingkat kesulitan yang sama, dan soal yang berbeda (Ary et al., 2010). Skor tersebut kemudian dikorelasikan untuk mengukur koefisien reliabilitas (James, 2013).

3. Reliabilitas sebagai konsistensi internal

Reliabilitas sebagai konsistensi internal menguji homogenitas item dalam suatu instrumen (Ary et al., 2010). Konsistensi internal suatu tes ditentukan dari satu kali pemberian tes (Rosaroso, 2015).

Seperti dikutip dalam Tosuncuoglu (2018), Heaton menyatakan bahwa beberapa faktor mempengaruhi reliabilitas tes.

1. Banyaknya sampel tes yang dipilih untuk pengujian

Ketika tes memiliki jumlah soal lebih banyak, tes akan lebih reliabel.

2. Fluktuasi dalam administrasi tes.

Reliabilitas tes bisa rendah jika kondisi tes cenderung berfluktuasi dari satu pelaksanaan ke pelaksanaan lainnya.

3. Faktor pribadi

Faktor pribadi berkaitan dengan kondisi fisik dan psikologis seperti kesehatan yang buruk, kelelahan, kurangnya minat atau motivasi, kecemasan, dan kesedihan.

4. Instruksi/petunjuk pengujian

Ketika instruksi/petunjuk tes jelas, hasil tes akan lebih reliabel.

5. Fluktuasi dalam penilaian

Subjektivitas dalam penilaian dapat menimbulkan inkonsistensi dalam skor dan menghasilkan pengukuran yang tidak dapat diandalkan/ tidak reliabel.

Ary et al. (2010) menyatakan bahwa faktor-faktor yang mempengaruhi reliabilitas adalah sebagai berikut.

1. Panjang tes

Ketika jumlah soal tes lebih banyak, skor sebenarnya lebih representatif.

2. Heterogenitas kelompok

Koefisien reliabilitas akan lebih tinggi bila mahasiswa yang mengikuti tes heterogen.

3. Kemampuan individu

Ketika mahasiswa memiliki kemampuan yang lebih tinggi, tes tersebut akan dapat diandalkan, tetapi tidak akan dapat diandalkan ketika mahasiswa memiliki kemampuan yang lebih rendah. Tingkat kesulitan tes juga mempengaruhi reliabilitas tes.

4. Teknik khusus yang digunakan untuk estimasi reliabilitas

Bentuk alternatif dengan teknik selang waktu memberikan estimasi reliabilitas yang lebih rendah daripada prosedur test-retest atau split-half.

5. Sifat variabel yang diukur

Tes prestasi akademik memiliki reliabilitas yang sangat tinggi (koefisien 0,90 atau lebih tinggi). Tes bakat memiliki reliabilitas yang lebih rendah (0,80 atau lebih rendah). Tes kepribadian memiliki reliabilitas sedang (0,60 hingga 0,70).

6. Objektivitas penilaian

Inkonsistensi penilaian mengurangi reliabilitas tes.

Brown (2004) menyatakan bahwa reliabilitas suatu tes dipengaruhi oleh beberapa faktor.

1. Reliabilitas Terkait Mahasiswa

Reliabilitas terkait mahasiswa yang paling umum

adalah karena faktor fisik atau psikologi. Ketika seorang mahasiswa mengambil dua tes dan lelah, tes tidak dapat diandalkan.

2. Reliabilitas Penilai

Reliabilitas antar-penilai terjadi ketika dua atau lebih skor menghasilkan skor yang tidak konsisten dari tes yang sama karena kurangnya perhatian pada kriteria penilaian dan kurangnya pengalaman.

3. Reliabilitas Administrasi/ Pelaksanaan Tes

Jika tes dikelola dengan baik, tes akan dapat diandalkan.

4. Reliabilitas Tes

Tes sendiri dapat menyebabkan kesalahan pengukuran.

Jenis tes adalah tes formatif dan tes sumatif. Tes formatif adalah evaluasi dan analisis kegiatan belajar mahasiswa sehari-hari. Ini membantu pendidik membimbing pembelajaran mahasiswa (Qu & Zhang, 2013). Tes sumatif dilakukan setelah semua materi dipelajari dan diselesaikan pada akhir semester. Dalam tes sumatif, pendidik dapat mendeteksi apa yang dapat diingat peserta didik tentang materi untuk diberikan nilai (Qu & Zhang dalam Sugianto, 2017). Dosen perlu

melakukan tes sumatif untuk mengukur kemampuan mahasiswa dalam menguasai semua materi dalam satu semester.

Di universitas, tes formatif dilakukan untuk semua mata kuliah, termasuk *grammar* (tata Bahasa). Berdasarkan Kurikulum KKNI (Kurikulum Berbasis Kerangka Kualifikasi Nasional Indonesia), *grammar* (tata bahasa) merupakan mata kuliah yang wajib dipelajari dan dikuasai oleh mahasiswa program studi pendidikan bahasa Inggris di universitas. Tata bahasa membutuhkan perhatian lebih dari dosen dalam pembelajaran bahasa Inggris (Calanoga, 2019). Tata bahasa membantu mahasiswa untuk menyusun kata menjadi kalimat gramatikal, yang melibatkan beberapa unit linguistik (Afdaleni, 2018; Richards dalam Sioco & Vera, 2018). Tata bahasa memudahkan mahasiswa untuk menyampaikan ide-ide mereka dengan jelas (Bradshaw dalam Sioco & Vera, 2018). Kompetensi tata bahasa yang baik memungkinkan mahasiswa untuk berkomunikasi secara efektif karena akurasi tata bahasa mereka (Marashi, 2014). Tanpa pengetahuan tata bahasa yang memadai, mahasiswa tidak

tahu bagaimana menyusun kalimat dengan benar dan memahami artinya (Cam & Tran, 2017).

Dalam penelitian ini, peneliti mengambil tiga penelitian yang relevan. Penelitian pertama dilakukan oleh Setiawaty et al. (2017) tentang validitas dan reliabilitas tes pilihan ganda bahasa Indonesia pada ujian akhir. Hasil perhitungan indeks reliabilitas adalah $0,3657 < 0,6$. Artinya, tes tidak dapat digunakan. Penelitian kedua dilakukan oleh Jayanti et al. (2019) tentang validitas dan reliabilitas ujian nasional bahasa Inggris. Hasil penelitian menunjukkan bahwa tes memenuhi kriteria validitas dan reliabilitas. Penelitian ketiga dilakukan oleh Setiabudi et al. (2019) tentang validitas dan reliabilitas tes buatan guru di sekolah menengah atas. Hasil analisis menunjukkan bahwa tes tersebut valid dan reliabel, namun keduanya berada pada kategori sedang.

Kesamaan penelitian ini dengan penelitian sebelumnya adalah semua penelitian menganalisis reliabilitas tes. Perbedaan penelitian ini dengan penelitian sebelumnya adalah penelitian pertama menganalisis validitas dan reliabilitas tes

pilihan ganda bahasa Indonesia pada ujian akhir, penelitian kedua menganalisis validitas dan reliabilitas ujian nasional bahasa Inggris, penelitian ketiga menganalisis validitas dan reliabilitas tes buatan guru di SMA, sedangkan penelitian ini menganalisis faktor-faktor yang mempengaruhi reliabilitas tes tata bahasa di universitas.

Dalam penelitian ini, peneliti bermaksud untuk menganalisis faktor-faktor yang mempengaruhi reliabilitas tes tata Bahasa Inggris di prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo, Indonesia.

Berdasarkan latar belakang penelitian, peneliti memiliki dua rumusan masalah:

1. Berapa reliabilitas tes tata bahasa Inggris di universitas di Indonesia?
2. Faktor-faktor apa saja yang mempengaruhi reliabilitas tes tata bahasa Inggris di universitas di Indonesia?

Sehubungan dengan rumusan masalah, tujuan penelitian adalah sebagai berikut:

1. Untuk menganalisis reliabilitas tes tata bahasa Inggris di universitas di Indonesia.
2. Menganalisis faktor-faktor yang mempengaruhi reliabilitas tes

tata bahasa Inggris di universitas di Indonesia.

METODE

Penelitian ini bersifat kualitatif (Creswell, 2012). Penelitian ini menggunakan studi kasus deskriptif. Hal ini bertujuan untuk menyajikan informasi rinci tentang fenomena tertentu untuk mendapatkan pemahaman yang mendalam tentang kasus tersebut (Heigham & Croker, 2009). Dalam penelitian ini, para peneliti menyajikan informasi rinci tentang faktor-faktor yang mempengaruhi reliabilitas tes tata Bahasa Inggris. Objek penelitian adalah 40 soal pilihan ganda tes formatif tata bahasa Inggris di prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo, Indonesia. Tes formatif tata bahasa dilakukan pada tanggal 30 September 2021 melalui *google form*. Subyek penelitian adalah 30 mahasiswa semester lima.

Instrumen penelitian ini adalah tes, angket tertutup, dan wawancara terbuka. Tes berbentuk pilihan ganda disajikan dengan *google form* dengan materi *English grammar* (tata bahasa Inggris) tentang *adjective clause*. Dalam melaksanakan tes, peneliti meminta izin kepada dosen dan mahasiswa prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo, Indonesia untuk mengambil data, membagikan link

google form untuk tes, menjelaskan kepada mahasiswa tentang cara mengerjakan tes, meminta mahasiswa untuk mengerjakan tes, meminta mahasiswa untuk mengirimkan hasil tes, dan mendownload hasil tes.

Dalam menganalisis reliabilitas tes tata bahasa Inggris, peneliti membaca hasil tes, mengolah data di SPSS, menghitung hasil reliabilitas dengan menggunakan Cronbach's Alpha, mengkategorikan hasil reliabilitas, menjelaskan hasil reliabilitas, dan memberi kesimpulan.

Peneliti juga menggunakan angket tertutup dan wawancara terbuka untuk mengetahui faktor-faktor yang mempengaruhi reliabilitas tes tata bahasa Inggris di prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo, Indonesia. Dalam membagikan angket tertutup, peneliti meminta izin kepada dosen dan mahasiswa prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo untuk mengambil data, membagikan link *google form* untuk pengisian angket, menjelaskan kepada mahasiswa tentang cara mengisi angket, meminta mahasiswa untuk mengisi angket, meminta mahasiswa untuk mengirim angket, dan mendownload hasil angket.

Dalam menganalisis angket, peneliti membaca hasil angket,

mengidentifikasi, mengkategorikan dan menjelaskan faktor-faktor yang mempengaruhi reliabilitas tes tata bahasa Inggris, dan menyimpulkan hasil.

Dalam melaksanakan wawancara terbuka, peneliti meminta izin kepada dosen dan mahasiswa prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo untuk mengambil data, menyiapkan pertanyaan wawancara, membagikan link zoom untuk wawancara, menjelaskan kepada mahasiswa untuk prosedur wawancara, mewawancarai mahasiswa melalui zoom dan merekamnya, dan menyimpan hasil wawancara.

Dalam menganalisis wawancara terbuka, peneliti mendengarkan hasil wawancara, mentranskrip hasil wawancara, mengidentifikasi, mengkategorikan dan menjelaskan faktor-faktor yang mempengaruhi reliabilitas tes, dan memberi kesimpulan.

HASIL PENELITIAN DAN PEMBAHASAN

Pada bagian ini, peneliti akan menyajikan temuan tentang reliabilitas tes *English grammar* (tata Bahasa Inggris) di prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo, Indonesia. Dalam penelitian ini, peneliti memberikan tes dua kali

untuk menganalisis reliabilitas tes. Para peneliti juga memberikan angket tertutup dan melakukan wawancara terbuka untuk menganalisis faktor-faktor yang mempengaruhi reliabilitas tes tata bahasa Inggris di prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo, Indonesia.

Hasil Reliabilitas Tes Tata Bahasa Inggris

Scale: ALL VARIABLES

Case Processing Summary

| | N | % |
|----------------|----|-------|
| Valid | 30 | 100.0 |
| Cases Excluded | 0 | .0 |
| Total | 30 | 100.0 |

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

| Cronbach's Alpha | N of Items |
|------------------|------------|
| .754 | 2 |

Item-Total Statistics

| | Scale Mean if Deleted | Scale Variance if Deleted | Corrected Item-Total Correlation | Cronbach's Alpha if Deleted |
|--------|-----------------------|---------------------------|----------------------------------|-----------------------------|
| test 1 | 88.2667 | 104.754 | .639 | . |
| test 2 | 81.2000 | 202.097 | .639 | . |

Sujarweni (2014) menyatakan bahwa tes itu reliabel jika nilai Cronbach 's Alpha > 0.6. Dalam tes ini, nilai Cronbach's Alpha adalah 0.754 > 0.6. Maka tes ini reliable dan dapat digunakan.

Hasil Angket Tertutup

Tabel 1. Faktor-Faktor yang Mempengaruhi Reliabilitas Tes

| N o. | Pernyataan | Sangat Tidak Setuju (%) | Tidak Setuju (%) | Setuju (%) | Sangat Setuju (%) |
|------|---|-------------------------|------------------|------------|-------------------|
| | Reliabilitas Terkait Mahasiswa | | | | |
| 1 | Ketika saya sakit, saya tidak bisa mengerjakan tes dengan baik. | 0 | 13.3 | 53.3 | 33.3 |
| 2 | Ketika saya lelah, saya tidak bisa mengerjakan tes dengan baik. | 3.3 | 16.7 | 63.3 | 16.7 |
| 3 | Ketika saya sedih, saya tidak bisa mengerjakan tes dengan baik. | 10 | 33.3 | 43.3 | 13.3 |
| 4 | Ketika saya | 0 | 20 | 66.7 | 13.3 |

| | | | | | |
|---|--|-----|------|------|------|
| | cemas, saya tidak bisa mengerjakan tes dengan baik. | | | | |
| 5 | Ketika saya marah, saya tidak bisa mengerjakan tes dengan baik. | 6.7 | 43.3 | 36.7 | 13.3 |
| | Reliabilitas Penilai | | | | |
| 6 | Ketika dosen memberikan instruksi yang tidak jelas dalam tes, itu akan mempengaruhi skor saya. | 3.3 | 3.3 | 56.7 | 36.7 |
| 7 | Ketika dosen membedakan mahasiswa, itu akan mempengaruhi skor saya. | 3.3 | 20 | 56.7 | 20 |
| 8 | Ketika dosen tidak menyukai saya, itu akan mempengaruhi skor saya. | 3.3 | 0 | 40 | 56.7 |
| 9 | Ketika dosen memberikan kriteria | 3.3 | 16.7 | 66.7 | 13.3 |

| | | | | | |
|----|--|-----|------|------|------|
| | skor yang tidak jelas, itu akan mempengaruhi skor saya. | | | | |
| 10 | Ketika dosen memberikan pertanyaan ambigu, itu akan mempengaruhi skor saya. | 0 | 10 | 63.3 | 26.7 |
| | Reliabilitas Pelaksanaan Tes | | | | |
| 11 | Saya tidak bisa mengerjakan tes dengan baik ketika jumlah soal terlalu banyak. | 0 | 33.3 | 50 | 16.7 |
| 12 | Saya tidak bisa mengerjakan tes dengan baik ketika cuaca terlalu panas. | 40 | 0 | 46.7 | 13.3 |
| 13 | Saya tidak bisa mengerjakan tes dengan baik ketika akses internet tidak | 3.3 | 0 | 40 | 56.7 |

| | | | | | |
|----|--|-----|------|------|------|
| | bagus. | | | | |
| 14 | Saya tidak bisa mengerjakan tes dengan baik ketika terlalu berisik. | 0 | 3.3 | 60 | 36.7 |
| 15 | Saya tidak bisa mengerjakan tes dengan baik ketika tempatnya tidak nyaman. | 3.3 | 3.3 | 63.3 | 30 |
| | Reliabilitas Tes | | | | |
| 16 | Saya tidak bisa mengerjakan tes dengan baik ketika tes terlalu lama. | 3.3 | 30 | 50 | 16.7 |
| 17 | Saya tidak bisa mengerjakan tes dengan baik ketika satu soal memiliki lebih dari satu jawaban. | 3.3 | 36.7 | 60 | 0 |
| 18 | Saya tidak bisa mengerjakan tes dengan baik dengan waktu yang | 0 | 10 | 46.7 | 43.3 |

| | | | | | |
|----|--|------|------|------|------|
| | sangat terbatas. | | | | |
| 19 | Saya tidak bisa mengerjakan tes dengan baik ketika ukuran font terlalu kecil. | 10 | 23.3 | 53.3 | 13.3 |
| 20 | Saya tidak bisa mengerjakan tes dengan baik ketika latar belakang google form gelap (tes online) atau ketika kualitas tinta printer kurang bagus (tes kertas). | 13.3 | 20 | 50 | 16.7 |

Pada pernyataan pertama, tidak ada mahasiswa yang sangat tidak setuju, 13,3% mahasiswa tidak setuju, 53,3% mahasiswa setuju, dan 33,3% mahasiswa sangat setuju bahwa ketika sakit, mereka tidak bisa mengerjakan tes dengan baik. Pada pernyataan kedua, 3,3% mahasiswa sangat tidak setuju, 16,7% mahasiswa tidak setuju, 63,3% mahasiswa setuju, dan 16,7% mahasiswa sangat setuju bahwa

ketika lelah, mereka tidak bisa mengerjakan tes dengan baik. Pada pernyataan ketiga, 10% mahasiswa sangat tidak setuju, 33,3% mahasiswa tidak setuju, 43,3% mahasiswa setuju, dan 13,3% mahasiswa sangat setuju bahwa ketika sedih, mereka tidak bisa mengerjakan tes dengan baik. Pada pernyataan keempat, tidak ada mahasiswa yang sangat tidak setuju, 20% mahasiswa tidak setuju, 66,7% mahasiswa setuju, dan 13,3% mahasiswa sangat setuju bahwa ketika cemas, mereka tidak bisa mengerjakan tes dengan baik. Pada pernyataan kelima, 6,7% mahasiswa sangat tidak setuju, 43,3% mahasiswa tidak setuju, 36,7% mahasiswa setuju, dan 13,3% mahasiswa sangat setuju bahwa ketika marah, mereka tidak bisa mengerjakan tes dengan baik. Pada pernyataan keenam, 3,3% mahasiswa sangat tidak setuju, 3,3% mahasiswa tidak setuju, 56,7% mahasiswa setuju, dan 36,7% mahasiswa sangat setuju bahwa ketika dosen memberikan instruksi yang tidak jelas dalam tes, itu akan mempengaruhi skor mereka. Pada pernyataan ketujuh, 3,3% mahasiswa sangat tidak setuju, 20% mahasiswa tidak setuju, 56,7% mahasiswa setuju, dan 20% mahasiswa sangat setuju bahwa ketika dosen membedakan mahasiswa, itu akan mempengaruhi

skor mereka. Pada pernyataan kedelapan, 3,3% mahasiswa sangat tidak setuju, tidak ada mahasiswa yang tidak setuju, 40% mahasiswa setuju, dan 56,7% mahasiswa sangat setuju bahwa ketika dosen tidak menyukai mereka, itu akan mempengaruhi skor mereka. Pada pernyataan kesembilan, 3,3% mahasiswa sangat tidak setuju, 16,7% mahasiswa tidak setuju, 66,7% mahasiswa setuju, dan 13,3% mahasiswa sangat setuju bahwa ketika dosen memberikan kriteria skor yang tidak jelas, itu akan mempengaruhi skor mereka. Pada pernyataan kesepuluh, tidak ada mahasiswa yang sangat tidak setuju, 10% mahasiswa tidak setuju, 63,3% mahasiswa setuju, dan 26,7% mahasiswa sangat setuju bahwa ketika dosen memberikan pertanyaan ambigu, itu akan mempengaruhi skor mereka.

Pada pernyataan kesebelas, tidak ada mahasiswa yang sangat tidak setuju, 33,3% mahasiswa tidak setuju, 50% mahasiswa setuju, dan 16,7% mahasiswa sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik ketika jumlah soal terlalu banyak. Pada pernyataan kedua belas, 40% mahasiswa sangat tidak setuju, tidak ada mahasiswa yang tidak setuju, 46,7% mahasiswa setuju, dan 13,3% mahasiswa sangat setuju bahwa mereka tidak bisa mengerjakan tes

dengan baik ketika cuaca terlalu panas. Pada pernyataan ketiga belas, 3,3% mahasiswa sangat tidak setuju, tidak ada mahasiswa yang tidak setuju, 40% mahasiswa setuju, dan 56,7% mahasiswa sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik ketika akses internet tidak bagus. Pada pernyataan keempat belas, tidak ada mahasiswa yang sangat tidak setuju, 3,3% mahasiswa tidak setuju, 60% mahasiswa setuju, dan 36,7% sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik ketika terlalu berisik. Pada pernyataan kelima belas, 3,3% mahasiswa sangat tidak setuju, 33,3% mahasiswa tidak setuju, 63,3% mahasiswa setuju, dan 30% sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik ketika tempatnya tidak nyaman. Pada pernyataan keenam belas, 3,3% mahasiswa sangat tidak setuju, 30% mahasiswa tidak setuju, 50% mahasiswa setuju, dan 16,7% mahasiswa sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik ketika tes terlalu lama. Pada pernyataan ketujuh belas, 3,3% mahasiswa sangat tidak setuju, 36,7% mahasiswa tidak setuju, 60% mahasiswa setuju, dan tidak ada mahasiswa yang sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik ketika satu soal memiliki lebih dari satu

jawaban. Pada pernyataan kedelapan belas, tidak ada mahasiswa yang sangat tidak setuju, 10% mahasiswa tidak setuju, 46,7% mahasiswa setuju, dan 43,3% mahasiswa sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik dengan waktu yang sangat terbatas. Pada pernyataan kesembilan belas, 10% mahasiswa sangat tidak setuju, 23,3% mahasiswa tidak setuju, 53,3% mahasiswa setuju, dan 13,3% mahasiswa sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik ketika ukuran font terlalu kecil. Pada pernyataan kedua puluh, 13,3% mahasiswa sangat tidak setuju, 20% mahasiswa tidak setuju, 50% mahasiswa setuju, dan 16,7% mahasiswa sangat setuju bahwa mereka tidak bisa mengerjakan tes dengan baik ketika latar belakang google form gelap (tes online) atau ketika kualitas tinta printer kurang bagus (tes kertas).

Berdasarkan hasil angket tertutup, faktor-faktor yang mempengaruhi reliabilitas tes adalah reliabilitas terkait mahasiswa, reliabilitas penilai, reliabilitas administrasi tes, dan reliabilitas tes.

Hasil Wawancara

Di sini, peneliti menjelaskan hasil wawancara terbuka kepada enam mahasiswa semester 5 jurusan pendidikan bahasa Inggris di

Universitas Muhammadiyah Purworejo, Indonesia.

1. Mahasiswa yang mendapat nilai tinggi dalam tes menyatakan telah mempelajari dan memahami materi dengan baik. Mahasiswa yang mendapat nilai sedang pada tes menyatakan bahwa mereka tidak memahami satu atau beberapa bagian materi. Mahasiswa yang mendapat nilai rendah pada tes menyatakan tidak mempelajari materi dan tidak memahami materi.
2. Kesulitan mahasiswa pada saat mengerjakan tes karena tidak menguasai materi pembelajaran yang diajarkan oleh dosennya.
3. Mahasiswa perlu belajar dengan baik untuk mendapatkan nilai yang baik.
4. Efek belajar sebelum tes adalah dapat memahami materi, mendapatkan nilai yang baik, dan merasa lebih percaya diri.
5. Kondisi fisik mahasiswa baik. Semua sehat.
6. Mahasiswa yang mendapat nilai tinggi memahami materi dengan baik. Mahasiswa yang mendapat nilai sedang tidak memahami satu atau beberapa bagian materi. Mahasiswa yang mendapat nilai rendah tidak memahami materi karena tidak mempelajarinya.

7. Mahasiswa menyatakan bahwa kejelasan instruksi/petunjuk tes akan berpengaruh signifikan terhadap hasil tes. Jika petunjuk tes tidak jelas, mereka tidak bisa mengerjakan tes dan mendapatkan nilai jelek.
8. Mahasiswa menyatakan bahwa efek dari kejelasan petunjuk adalah mereka dapat memahami tes, mengerjakan tes lebih cepat, dan mendapatkan nilai yang lebih baik.
9. Mahasiswa menyatakan bahwa kriteria penilaian baik dan sangat jelas. Mereka menyatakan bahwa kriteria penilaian yang tidak jelas akan mempengaruhi skor mereka karena inkonsistensi penilaian.
10. Mahasiswa menyatakan bahwa suasana saat mengerjakan tes tenang dan nyaman.
11. Mahasiswa yang mendapat nilai tinggi dan sedang menyatakan memiliki waktu yang cukup untuk mengerjakan tes, sehingga mendapat nilai baik. Namun, mahasiswa yang mendapat nilai rendah menyatakan waktunya tidak mencukupi. Ada pengaruh antara durasi tes terhadap hasil tes mahasiswa.
12. Mahasiswa menyatakan bahwa tidak ada pertanyaan yang membingungkan dalam tes.

SIMPULAN

Berdasarkan hasil penelitian, reliabilitas tes tata bahasa Inggris di prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo, Indonesia adalah 0,754, sehingga tes tersebut reliabel. Dilihat dari hasil angket tertutup, faktor-faktor yang mempengaruhi reliabilitas tes tata bahasa Inggris di prodi Pendidikan Bahasa Inggris Universitas Muhammadiyah Purworejo, Indonesia adalah reliabilitas terkait mahasiswa, reliabilitas penilai, reliabilitas pelaksanaan tes, dan reliabilitas tes. Dilihat dari hasil wawancara, faktor-faktor yang mempengaruhi reliabilitas tes adalah penguasaan materi pembelajaran, persiapan belajar sebelum tes, kondisi fisik mahasiswa, kejelasan instruksi/petunjuk tes, kejelasan soal dalam tes, suasana dan alokasi waktu untuk mengerjakan tes.

DAFTAR PUSTAKA

- Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and*

- Research in Education*, 9(1), 109–119.
<https://doi.org/10.11591/ijer.e.v9i1.20457>
- Afdaleni. (2018). Students' Interest Level in Learning English Grammar. *Asian EFL Journal*, 20(4), 61–65.
- Apsari, Y., & Haryudin, A. (2017). The Analysis of English Lecturers' Classroom-Based Reading Assessments to Improve Students' Reading Comprehension. *ELTIN Journal*, 5(1), 35–44.
- Arifin, M. A. (2018). Validity, Reliability and Practicality Of The First Certification in English (FCE) and The Business Language Testing Service (Bulats). *Journal of Language Teaching and Learning, Linguistics and Literature*, 6(2), 80–95.
- Ary, D., Jacobs, L. C., Sorensen, C., & Razavieh, A. (2010). *Introduction to Research in Education* (8th ed.). Wadsworth Cengage Learning.
- Azmi, U. (2020). Developing web-based reading tests for the students of English language education. *Journal of Applied Linguistics, Translation, and Literature*, 1(2), 92–104.
- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. Longman.
- Calanoga, M. C. M. (2019). Mechanics and Grammar Error Analysis in Students' Write-Ups: Basis for Incidental Teaching in the Classroom. *Asian EFL Journal*, 23(3.3), 104–125.
- Cam, L., & Tran, T. M. T. (2017). An Evaluation of using Games in Teaching English Grammar for First Year English-Majored Students at Dong Nai Technology University. *International Journal of Learning, Teaching and Educational Research*, 16(7), 55–71.
- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (Fourth). Pearson Education.
- Furwana, D. (2019). Validity and Reliability of Teacher-Made English Summative Test at Second Grade of Vocational High School 2 Palopo. *LANGUAGE CIRCLE: Journal of Language and Literature*, 13(2).
<http://journal.unnes.ac.id>
- Heigham, J., & Croker, R. A. (2009). *Qualitative Research in Applied Linguistics: A*

- Practical Introduction*.
Palgrave Macmillan.
- James, C. (2013). *Errors in Language Learning and Use: Exploring Error Analysis*. Routledge.
- Jayanti, D., Husna, N., & Hidayat, D. N. (2019). The Validity and Reliability Analysis of English National Final Examination for Junior High School. *Voices of English Language Education Society*, 3(2), 127–135.
- Marashi, H. (2014). Using Inductive and Deductive Consciousness Raising Tasks to Improve Field-Dependent and Field-Independent EFL Learners' Grammar. *The Iranian EFL Journal*, 10(3), 60–74.
- Öz, H., & Özturan, T. (2018). Computer-based and paper-based testing: Does the test administration mode influence the reliability and validity of achievement tests? *Journal of Language and Linguistic Studies*, 14(1), 67–85.
- Qu, W., & Zhang, C. (2013). The Analysis of Summative Assessment and Formative Assessment and Their Roles in College English Assessment System. *Journal of Language Teaching and Research*, 4(2), 335–339.
- <https://doi.org/10.4304/jltr.4.2.335-339>
- Rosaroso, R. C. (2015). Using Reliability Measures in Test Validation. *European Scientific Journal*, 11(18), 369–377.
- Setiabudi, A., Mulyadi, & Puspita, H. (2019). An Analysis of Validity and Reliability of A Teacher-Made Test (Case Study at XI Grade of SMA N 6 Bengkulu). *Journal of English Education and Teaching*, 3(4), 522–532.
- Setiawaty, R., Sulistyorini, T. B., Margono, & Rahmawati, L. E. (2017). Validity Test and Reliability of Indonesian Language Multiple Choice in Final Term Examination. *The 1st International Seminar on Language, Literature and Education, 2018*, 43–50. <https://doi.org/10.18502/kss.v3i9.2609>
- Sioco, E. C., & Vera, Dr. P. V. D. (2018). Grammatical Competence of Junior High School Students. *TESOL International Journal*, 13(2), 82–94.
- Sugianto, A. (2017). Validity and Reliability of English Summative Test for Senior High School. *Indonesian EFL Journal: Journal of ELT*,

- Linguistics, and Literature*, 3(2), 22–38.
- Sultana, R. (2015). Reliability of the Currently Administered Language Tests in Bangladesh: A Case Study. *Journal of Literature, Languages and Linguistics*, 17, 76–85.
- Taherdoost, H. (2016). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research in Management*, 5(3), 28–36.
- Tosuncuoglu, I. (2018). Importance of Assessment in ELT. *Journal of Education and Training Studies*, 6(9), 163–167. <https://doi.org/10.11114/jets.v6i9.3443>
- Ulfah, A. A., Kartono, & Susilaningsih, E. (2020). Validity of Content and Reliability of Inter-Rater Instruments Assessing Ability of Problem Solving. *Journal of Educational Research and Evaluation*, 9(1), 1–7.
- Wenno, I. H., Tuhurima, D., & Manoppo, Y. (2021). How to Create a Good Test. *Jurnal Pendidikan Profesi Guru Indonesia*, 1(1), 11–20.
- Zimmerman, D. W., & Zumbo, B. D. (2015). Resolving the Issue of How Reliability is Related to Statistical Power: Adhering to Mathematical Definitions. *Journal of Modern Applied Statistical Methods*, 14(2), 9–26.