

ABSTRACT

Icha Anggreaini Pratama, 2321044, An Analysis of Listening Test Items at the 9th Grade of SMP N 1 Bukittinggi, A Thesis, State Islamic University of Sjech M. Djamil Djambek Bukittinggi, 2025.

This research is conducted due to several problems in the English listening test used by the teacher. The English teacher did not consider the difficulty level of the test items, did not analyze whether the items could distinguish students' abilities, did not conduct item analysis, and used a test with an unbalanced distribution of difficulty levels. Therefore, this research aimed to analyze the difficulty level and discrimination index of English listening test items.

This research used a descriptive quantitative method. The data were obtained from students' scores on an English listening test consisting of 10 true/false items. The data were analyzed using formulas proposed by Brown and Abeywickrama to determine the difficulty level and discrimination index of the test items.

The results showed that the listening test items made by the teacher were valid but not reliable, with a reliability coefficient of 0.40. Based on the difficulty level analysis, 6 out of 10 items (60%) were categorized as easy and 4 items (40%) were categorized as moderate. No difficult items were found, indicating that the test was relatively easy for the students. The discrimination index analysis showed that 5 items (50%) were classified as good, 4 items (40%) as satisfactory, and 1 item (10%) as poor. These results indicate that most of the test items were not able to effectively distinguish between students with high and low ability. The low reliability of the test may have been caused by several factors, such as noisy testing environment, the absence of test reliability analysis before administration, and the small number of test items. In conclusion, the English listening test constructed by the teacher was not reliable and therefore could not be considered a good test. Teachers are expected to pay attention to the characteristics of a good test, including validity, reliability, and practicality, as well as item characteristics such as difficulty level and discrimination index when developing test instruments.

Keywords: *Item Analysis, Listening Test, Difficulty Level, Discrimination Index*

ABSTRAK

Icha Anggreaini Pratama, 2321044, Analisis Butir Soal Kemampuan Menyimak pada Siswa Kelas 9 di SMP N 1 Bukittinggi, Universitas Islam Negeri Sjech M. Djamil Djambek Bukittinggi, 2025.

Penelitian ini dilakukan karena adanya beberapa permasalahan pada tes menyimak Bahasa Inggris yang digunakan oleh guru. Guru tidak memperhatikan Tingkat kesulitan butir soal, tidak menganalisis apakah butir soal dapat membedakan kemampuan siswa, tidak melakukan analisis butir soal, serta menggunakan tes dengan distribusi Tingkat kesulitan yang tidak seimbang. Oleh karena itu, penelitian ini bertujuan untuk menganalisis Tingkat kesulitan dan daya pembeda butir soal tes menyimak Bahasa Inggris.

Penelitian ini menggunakan metode deskriptif kuantitatif. Data penelitian diperoleh dari skor siswa pada tes menyimak Bahasa Inggris yang terdiri dari 10 butir soal benar/salah. Data dianalisis menggunakan rumus yang dikemukakan oleh Brown dan Abeywickrama untuk menentukan Tingkat kesulitan dan indeks daya pembeda butir soal.

Hasil penelitian menunjukkan bahwa butir soal menyimak yang dibuat oleh guru bersifat valid tetapi tidak reliable, dengan koefisien reliabilitas sebesar 0.40. Berdasarkan analisis Tingkat kesulitan, 6 dari 10 butir soal (60%) tergolong mudah dan 4 butir soal (40%) tergolong sedang. Tidak ditemukan butir soal yang tergolong sulit, sehingga dapat disimpulkan bahwa tes tersebut relatif mudah bagi siswa. Analisis daya pembeda menunjukkan bahwa 5 butir soal (50%) tergolong baik, 4 butir soal (40%) tergolong cukup, dan 1 butir soal (10%) tergolong kurang. Hasil ini menunjukkan bahwa Sebagian besar butir soal belum mampu membedakan secara efektif antara siswa berkemampuan tinggi dan rendah. Rendahnya reliabilitas tes kemungkinan disebabkan oleh beberapa faktor, seperti kondisi lingkungan yang bising saat pelaksanaan tes, tidak adanya analisis reliabilitas sebelum tes diberikan, serta jumlah butir soal yang terbatas. Kesimpulannya, tes menyimak Bahasa Inggris yang disusun oleh guru belum reliabel sehingga belum dapat dikategorikan sebagai tes yang baik. Guru diharapkan dapat memperhatikan karakteristik tes yang baik, yaitu validitas, reliabilitas, dan kepraktisan, serta Karakteristik butir soal seperti tingkat kesulitan dan daya pembeda dalam Menyusun instrument tes.

Kata Kunci: Analisis Butir Soal, Tes Menyimak, Tingkat Kesulitan, Daya Pembeda