

BAB I

PENDAHULUAN

Pada bab ini, peneliti menyajikan beberapa topik seperti latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup dan batasan penelitian, dan definisi istilah kunci penelitian.

A. Latar Belakang Penelitian

Undang-Undang Sisdiknas No. 20/2003, Pasal 58, Ayat 1 yang berbunyi "Evaluasi hasil belajar peserta didik dilakukan oleh pendidik untuk memantau proses, kemajuan, dan perbaikan hasil belajar peserta didik secara berkesinambungan." menuntut guru untuk mengevaluasi prestasi belajar siswa guna memantau proses, dan kemajuan, serta meningkatkan hasil belajar siswa. Artinya, dalam proses pembelajaran, tugas guru selain menyampaikan materi yaitu juga dengan mengukur kemampuan siswa. Guru berperan untuk mengevaluasi prestasi siswa sehingga hasil evaluasi dapat dijadikan pedoman bagi kualitas pengajaran yang lebih baik. Menurut Allen evaluasi adalah proses sistematis mengumpulkan dan menganalisis data untuk membuat penilaian tentang program atau kegiatan tertentu. Pada kegiatan ini, guru memerlukan pengumpulan, analisis, dan interpretasi informasi tentang belajar mengajar untuk meningkatkan prestasi akademik siswa dalam proses pendidikan mereka.¹ Salah satu instrumen evaluasi adalah tes. Sebuah tes, menurut Postlethwaite, adalah instrumen atau teknik yang

¹ Allen, D. (1998). *Assessing Student Learning*. New York: Teachers College Press.

menyajikan serangkaian tugas dimana siswa harus mengerjakan dan menyelesaikan tugas tersebut. Seorang guru dapat mengetahui kemampuan siswa melalui tes. Soal tes yang baik sangat penting untuk mengukur berhasil atau tidaknya proses belajar pembelajaran. Dengan demikian, guru dapat mengevaluasi kinerja siswa dengan baik apabila butir tes yang dibuat sudah terbukti bagus dan memenuhi kriteria tes yang berlaku.²

Menurut Hughes, kualitas tes dapat diukur melalui validitas, reliabilitas, dan kepraktisan. Validitas mengacu pada tes yang dapat benar-benar mengukur apa yang ingin diukur.³ Sementara itu, menyatakan reliabilitas sebagai konsistensi nilai ujian. Ini juga mengacu pada berbagai hasil tes siswa yang mana ketika diujikan kembali, mereka mendapatkan nilai yang tidak terlalu jauh berbeda dari nilai sebelumnya.⁴ Hal ini sejalan dengan Bachman dan Palmer, yang menyatakan bahwa reliabilitas adalah konsistensi nilai tes. Baik validitas maupun reliabilitas tidak dapat dipisahkan dari penilaian. Demikian pula, Kepraktisan, sesuai dengan Bachman dan Palmer, mendefinisikan sebagai hubungan antara sumber daya yang tersedia seperti sumber daya manusia, bahan, waktu, dan sumber daya untuk tes yang akan diperlukan dalam membuat kerangka, pengembangan, dan penggunaan tes.⁵ Dengan kata lain, kepraktisan berkaitan dengan tes yang mudah dirancang, dikelola, dan dinilai. Terlepas dari seberapa valid sebuah tes, tes tersebut tetaplah

² Postlethwaite, T. N. (2005). *Educational research: some basic concepts and terminology*. Quantitative research methods.

³ Hughes, A. (2005). *Testing for Language Teachers (2nd ed)*. London: Cambridge University Press.

⁴ Haryudin, A. (2015). *Validity and reliability of English summative tests at junior high school in West Bandung*. Jurnal Ilmiah P2P STKIP Siliwangi, 2(1), 77-90.
<https://doi.org/10.22460/p2m.v2i1p77-90.167>

⁵ Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice. Designing and developing useful language tests*. Oxford, New York: Oxford University Press.

harus memenuhi kriteria kepraktisannya.

Dengan adanya implementasi kurikulum merdeka yang dilaksanakan pada tahun ajaran baru, bentuk penilaian dibagi menjadi dua yakni penilaian formatif dan sumatif. Seperti yang diketahui, beberapa tahun lalu penilaian sumatif seperti Penilaian Tengah Semester (PTS) atau Penilaian Akhir Tahun (PAS) dilaksanakan dengan serentak yang mana konsep soal dibuat oleh tim MGMP per daerah telah dirubah. Pada penerapan kurikulum merdeka, penilaian dikembalikan ke guru masing-masing bidang pelajaran. Oleh karena itu, setidaknya dibutuhkan pengetahuan dasar untuk guru mengetahui bagaimana membuat soal yang baik untuk mengukur kemampuan siswa.

Penelitian di bidang penilaian dan evaluasi oleh Cang dan Wu yang mengevaluasi tentang validitas dan reliabilitas sebuah tes yang dibuat oleh guru dalam konteks sistem penilaian portofolio berbasis web. Mereka menyebutkan hasil penilaian portofolio berbasis web berikut ini:

1. Mencapai tingkat keandalan yang dapat diterima; menunjukkan tingkat keandalan antar penilai dan penilai dalam yang kuat, yang dianggap sebagai metode penilaian yang andal;
2. Membuktikan tingkat validitas yang dapat diterima;
3. Nilai portofolio sangat sesuai dengan nilai ujian akhir siswa, yang menyiratkan bahwa penilaian guru portofolio berbasis web adalah penilaian yang valid.⁶

⁶ Chang, C. C., Chen, C. C., & Chen, Y. H. (2012). *Reflective behaviors under a web-based portfolio assessment environment for high school students in a computer course*. *Computers and Education*, 58(1), 459–469. <https://doi.org/10.1016/j.compedu.2011.08.023>.

Demikian pula, Muhammad yang menganalisis perbandingan antara tes buatan untuk First Certification in English (FCE) dan Business Language Testing Service (BULATS) dalam hal validitas, reliabilitas, dan kepraktisan. Hasil penelitian tersebut menunjukkan bahwa BULATS memiliki hasil reliabilitas yang lebih tinggi daripada FCE. Sementara itu, baik BULATS maupun FCE memiliki validitas isi karena dibangun oleh sekelompok ahli pengujian bahasa. Dari segi kepraktisan, BULATS dan FCE bisa lebih praktis jika tidak mengharuskan penguji berada di ruang ujian.⁷

Di Indonesia juga ada beberapa penelitian yang berkaitan dengan evaluasi. Sugianto dan Jayanti, Husna, dan Hidayat menyelidiki Asesmen Nasional yang didistribusikan ke seluruh sekolah di Indonesia. Hasilnya menunjukkan bahwa tes tersebut memiliki validitas isi yang tinggi dan dianggap reliabel.⁸ ⁹ Furwana meneliti tentang validitas dan reliabilitas tes sumatif Bahasa Inggris yang dibuat oleh guru di sebuah sekolah menengah kejuruan di Palopo. Hasil penelitian menunjukkan bahwa tes sumatif bahasa Inggris memiliki validitas dan reliabilitas isi yang tinggi. Hasil penelitian menunjukkan bahwa tes buatan guru memiliki kualitas yang baik dari segi validitas dan reliabilitas isi.¹⁰ Sebaliknya, Primadani

⁷ Muhammad, A. A. (2018). *Validity, reliability and practicality of the first certification in English (FCE) and the Business Language Testing Service (BULATS)*. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 6(2). <https://doi.org/10.24256/ideas.v6i2.520>.

⁸ Sugianto, A. (2016). *An Analysis of English National Final Examination for Junior High School in Terms of Validity and Reliability*. *Journal on English as a Foreign Language*, 6(1), 31-42.

⁹ Jayanti, D., Husna, N., & Hidayat, D. N. (2019). *The validity and reliability analysis of English National Final Examination for Junior High School*. *VELES Voices of English Language Education Society*, 3(2), 127. <https://doi.org/10.29408/veles.v3i2.1551>.

¹⁰ Furwana, D. (2019). *Validity and reliability of teacher-made English summative test at second grade of vocational high school 2 Palopo*. *Journal of Language and Literature*, 13(2), 113-122. <https://doi.org/10.15294/lc.v13i2.18967>.

dan Ratnafuri menguji ujian tengah semester dan ujian akhir Bahasa Inggris. Kedua penelitian tersebut menunjukkan bahwa kualitas tes buruk karena guru tidak memenuhi syarat-syarat dalam merancang item tes, yang mana hal tersebut membuat kualitas tes menjadi lebih rendah.^{11 12} Namun, berbeda dengan penelitian lain yang dilakukan oleh Semiun dan Luruk mereka menganalisis kualitas tes sumatif bahasa Inggris yang dibuat oleh guru untuk siswa kelas 7 di Kupang NTT. Item tes yang dibuat oleh guru menunjukkan hasil yang cukup memuaskan di mana telah memenuhi kriteria tes yang bagus. Kesimpulannya, hasil dari penelitian-penelitian tersebut menunjukkan hasil yang bervariasi pada tes buatan guru karena keahlian pembuat tes juga terkait dengan mengonstruksi tes.¹³

Namun, berdasarkan pengamatan dari tes sebelumnya, peneliti menemukan banyak masalah, salah satunya adalah ditemukannya beberapa kesalahan pertanyaan seperti gambar tidak muncul karena tes dilakukan menggunakan Google Form tidak menggunakan kertas, pilihan yang ambigu karena ada lebih dari satu jawaban yang benar, dan sebagainya. Selain itu, tes tersebut juga dibagikan kepada siswa tanpa adanya uji coba sehingga tidak dapat menjamin tingkat reliabilitas tes tersebut. Oleh karena itu, perlu untuk diteliti dan ditinjau lebih lanjut apakah tes dibuat oleh guru pada siswa kelas XII MAN Kota Mojokerto sudah memenuhi standar yang ditentukan atau belum untuk mengukur kemampuan belajar siswa.

¹¹ Primadani, A. E. (2013). *An analysis of a midterm English test of the 7th grade accelerated class at SMPN 3 Malang*. State University of Malang.

¹² Ratnafuri, W. F. (2011). *An analysis of the teacher-made English test in the final test of the 2nd semester of 2010/2011 of the first year students of SMPN 1 Kauman, Tulungagung*. State University of Malang.

¹³ Semiun, T. T., & Luruk, F. D. (2020). *The quality of an English summative test of a public junior high school, Kupang-NTT*. *English Language Teaching Educational Journal*, 3(2), 133–141. <https://doi.org/10.12928/eltej.v3i2.2311>.

Dalam penelitian ini, peneliti berfokus pada Asesmen Madrasah mata pelajaran bahasa Inggris pada siswa kelas XII MAN Kota Mojokerto tahun ajaran 2022/2023. Peneliti memilih MAN Kota Mojokerto karena peneliti ingin melihat kualitas tes pelajaran bahasa Inggris yang dibuat oleh guru. Selain itu kemudahan akses menjadi alasan lain peneliti memilih sekolah ini sebagai subjek penelitian. Selain itu, peneliti juga dapat mengamati instruksi kelas dan kegiatan pembelajaran sesuai kebutuhan karena peneliti adalah salah satu guru di sana.

Para siswa kelas XII dipilih karena mereka berada di kelas senior sebelum melanjutkan ke tingkat yang lebih tinggi berikutnya dan juga siswa kelas XII nantinya akan menghadapi ujian sekolah yang mana soal tes juga dibuat oleh guru tersebut. Untuk mengetahui kualitas soal tes, analisis dapat dilakukan dengan melihat validitas, reliabilitas, tingkat kesulitan butir soal, daya pembeda butir soal, dan keefektifan pengecoh soal. Aspek-aspek tersebut dianalisis dengan menggunakan program ITEMAN 4.3 sebagai solusi yang efektif untuk analisis soal tes karena sederhana, efisien, namun dengan hasil yang menyeluruh.

B. Rumusan Masalah

Berdasarkan latar belakang penelitian, peneliti memfokuskan pada penelitian dan analisis Asesmen Madrasah mata pelajaran bahasa Inggris yang dirancang untuk kelas XII MAN Kota Mojokerto. Sejalan dengan itu, peneliti merumuskan masalah penelitian sebagai berikut:

1. Bagaimana kualitas butir soal Asesmen Madrasah mata pelajaran bahasa Inggris kelas XII MAN Kota Mojokerto ditinjau dari validitas, reliabilitas, tingkat kesulitan

butir soal, daya pembeda butir soal, dan keefektifan pengecoh butir soal?

C. Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk mengetahui kualitas butir soal Asesmen Madrasah mata pelajaran bahasa Inggris pada siswa kelas XII MAN Kota Mojokerto ditinjau dari validitas, reliabilitas, tingkat kesulitan butir, diskriminasi butir soal, dan keefektifan pengecoh butir soal.

D. Manfaat Penelitian

Ada dua macam signifikansi dalam penelitian ini yaitu signifikansi teoritis dan signifikansi praktis. Secara teoritis, hasil penelitian ini diharapkan dapat bermanfaat sebagai referensi terkait bagi peneliti selanjutnya yang tertarik untuk melakukan penelitian serupa atau untuk menambah, memperbaiki atau membantah teori yang lama.

Secara praktis, penelitian ini diharapkan dapat memberikan manfaat dan kontribusi bagi guru khususnya guru pengajar untuk membuat ujian yang baik di masa yang akan datang. Bagi para guru, penelitian ini diharapkan dapat membantu mereka untuk mengevaluasi dan mempersiapkan materi yang cocok untuk ujian selanjutnya. Selain itu juga untuk menyusun tes yang terkonstruksi dengan baik dengan memenuhi kriteria butir tes yang baik. Bagi Kepala Sekolah, hal ini bermanfaat sebagai pedoman untuk meningkatkan kualitas sekolah dalam menyediakan dan memfasilitasi tes bagi siswa.

E. Ruang Lingkup dan Batasan

Ruang lingkup penelitian ini adalah kualitas butir tes Asesmen Madrasah mata pelajaran Bahasa Inggris pada siswa kelas XII MAN Kota Mojokerto. Batasan penelitian ini adalah bahwa penelitian hanya dapat dilakukan di kelas XII dimana peneliti mendapat izin dan akses untuk melakukan penelitian.

F. Definisi Istilah Kunci

1. Tes: tes adalah suatu teknik untuk mengukur keterampilan, pengetahuan, atau kinerja seseorang dalam suatu bidang tertentu; namun demikian, membuat tes yang sukses adalah usaha keras yang melibatkan sains dan seni (Brown, 2004)
2. Analisis soal adalah prosedur sistematis yang dilakukan peneliti untuk mengungkap informasi tentang kualitas soal yang ditinjau dari validitas, reliabilitas, tingkat kesulitan, daya pembeda soal, dan pengecoh soal.
3. Asesmen Madrasah mata pelajaran Bahasa Inggris adalah tes bahasa Inggris yang digunakan sebagai ujian kenaikan kelas yang dibuat oleh guru bahasa Inggris pada siswa kelas XII di MAN Mojokerto tahun pelajaran 2022/2023 yang bertujuan untuk mengukur prestasi siswa.