

BAB I PENDAHULUAN

1.1 Latar Belakang

Pendidik meyakini bahwa penilaian di dalam kelas merupakan jembatan yang menghubungkan antara pengajaran dengan pembelajaran (Larbi, 2019). Bagaimana tidak, melalui serangkaian tes seorang pendidik mendapatkan jawaban apa yang harus diajarkan dan bagaimana mengajarkannya. Efek domino lainnya berkaitan dengan peningkatan proses pendidikan, motivasi belajar, yang kemudian berujung pada kualitas hasil belajar. Jimaa, (2011) menyatakan melalui tes akan membantu peserta didik memahami materi, informasi yang memuat progress belajar peserta didik, dan dasar pertimbangan keputusan bagi pendidik.

Berdasarkan tujuan, penilaian dikelompokkan menjadi sumatif dan formatif, tujuan penilaian sumatif meliputi seleksi, akuntabilitas, dan pemantauan yang efektif, sedangkan sumatif meliputi motivasi, umpan balik, dan peningkatan pembelajaran (Falchikov, 2005). Selaras studi Muhammad *et al.*, (2024) penilaian sumatif akan mengungkap pencapaian pembelajaran secara lebih luas. Lebih lanjut hasil tes sumatif memiliki proporsi yang besar dalam menentukan bilangan kuantitatif pada saat pelaporan, sehingga Larbi (2019) menyatakan hasil penilaian sumatif mempengaruhi karir pada jenjang yang lebih tinggi. Pengimplementasian tes sumatif yang didesain sesuai dengan tujuan pembelajaran merupakan aspek penting dalam praktik pendidikan. Jika instrumen tes sumatif dirancang dengan tepat hasil tes sumatif akan mengungkap data objektif untuk mendukung penilaian guru yang lebih profesional, menghasilkan keputusan tepat, dasar pertimbangan untuk penyesuaian kurikulum, dan meningkatkan proses pembelajaran (States *et al.*, 2018).

Ketepatan rancangan binstrumen tes berkaitan dengan parameter kualitas tes yang dapat ditentukan melalui analisis tes klasik yang meliputi reliabilitas, validitas, indeks kesukaran, dan terkhusus soal pilihan ganda adalah daya pembeda. Studi oleh Rismaulhijjah & Kuswanti (2022) hasil analisis butir menunjukkan validitas soal pada kategori valid 25% sedangkan 75% dinyatakan tidak valid, reliabilitas rendah, dan tidak HOTS (*high order thinking skill*). Selaras studi Fiska dkk., (2021) soal ulangan memiliki validitas rendah, tidak reliabel, dan daya pembeda kurang baik. Sejalan dengan studi Amri & Syahrani (2018) menemukan bahwa tes yang dibuat oleh guru belum diketahui daya beda dan tingkat kesukaran. Sebaliknya, hasil studi Hartono dkk., (2024) 93% soal valid, reliabilitas cukup, indeks kesulitan dan daya pembeda yang baik. Sehingga kualitas butir soal harus dilakukan peninjauan secara khusus.

Mahendra (2023) menyatakan bahwa Sekolah Menengah Pertama terbaik Jambi adalah SMP Swasta Xaverius 1 Jambi, didasarkan pada rata-rata nilai ujian nasional 2019. Fakta tersebut menunjukkan pada sekolah tersebut telah berlangsung proses-proses pembelajaran yang optimal, namun studi pendahuluan menunjukkan bahwa soal tes sumatif yang dibuat guru IPA SMP Xaverius 1 Jambi belum dilakukan analisis butir soal pada tes sumatif atau umumnya dikenal dengan istilah ujian akhir sekolah (UAS). Tentunya hal ini akan berimplikasi terhadap hasil pengukuran yang berpotensi bias. Selaras States *et al.*, (2018) jika soal sumatif yang dirancang tidak tepat, hasil penilaian hanya akan menyita waktu pengajaran, meningkatkan stres guru, dan hasil yang bias. Kemudian Oluwatayo (2012) menyatakan instrumen yang tidak valid atau tidak reliabel akan menghasilkan data yang menyebabkan regresi statistik dan mendistorsi hasil tes

sumatif. Serta soal dengan kualitas buruk mengandung sifat bias pada butir-butirnya (Retnawati & Hidayati, 2006).

Menurut Magdalena, dkk., (2021) analisis butir soal diperlukan untuk menilai kualitas soal. Sijabat, dkk., (2024) dalam studinya menyimpulkan bahwa soal atau instrumen tes yang telah dibuat setelah dianalisis dan menunjukkan hasil yang belum reliabel dan belum valid perlu dilakukan revisi. Pusat Penilaian Pendidikan (2013) menyatakan bahwa pertimbangan untuk mendapatkan hasil pengukuran yang baik, hasil dari pengukuran tersebut harus reliabel dan valid. Guru harus mempertimbangkan prinsip-prinsip dasar evaluasi dan menggunakan alat ukur yang valid, kredibel, dan praktis untuk meningkatkan kualitas evaluasi (Wijaya, dkk., 2024). Kriteria utama butir soal yang berkualitas adalah validitas, reliabilitas, indeks kesukaran, dan daya pembeda (Tuckman, 1985; Secolsky & Denison, 2017; Rahman & Nasyrah 2019; Inanna dkk., 2021; Wati, *et al.*, 2024).

Menurut Adom, *et al.*, (2020) dalam praktek pembelajaran guru melakukan penyesuaian sehingga memungkinkan evaluasi yang lebih spesifik dan personal. Melalui studi pendahuluan berupa wawancara kepada guru IPA di SMP Xaverius 1 Jambi diketahui bahwa soal pilihan ganda merupakan jenis soal yang dipilih untuk melakukan tes sumatif yang terintegrasi dengan ExamBrowser menggunakan Smartphone. Pengerjaan melalui ExamBroswer menjamin peserta didik dalam mengerjakan soal tidak membuka halaman lain. Pengerjaan menggunakan metode ini sangat efisien dikarenakan hasil tes yang lebih objektif, *paperless*, dan pengolahan hasil yang lebih singkat.

Kesadaran penuh akan pentingnya penilaian untuk mengungkap keberhasilan suatu proses pembelajaran, sehingga sangat perlu

mempertimbangkan kualitas tes yang diberikan agar menghasilkan data yang objektif. Objektivitas suatu tes sumatif bergantung pada validitas, reliabilitas, indeks kesukaran, dan daya pembeda. Mengingat fakta bahwa SMP Xaverius 1 Jambi belum melakukan uji tes klasik pada soal-soal UAS yang diberikan kepada peserta didik, sehingga sangat penting untuk dilakukan penelusuran untuk mengungkap kualitas tes yang diimplementasikan. Belum lagi, hasil studi literatur pada Lampiran 3 menunjukkan masih sangat jarang dilakukan penelitian di SMP Xaverius 1 Jambi.

Berlandaskan pada studi pendahuluan, studi literatur, wawancara, dan observasi yang telah dilakukan, maka tampak jelas bahwa kedudukan tes merupakan hal yang sangat krusial dalam sebuah pembelajaran. Namun pada prakteknya telah terjadi ketidaksesuaian dengan teori. Mengingat potensi analisis statistik dalam mendeterminasikan kualitas suatu tes, sehingga diajukan penelitian dengan judul *“Analisis Butir Soal Tes Sumatif yang Dibuat Guru Mata Pelajaran IPA di SMP Xaverius 1 Jambi”*. Melalui penelitian ini akan diidentifikasi kualitas soal ujian akhir yang diberikan kepada peserta didik.

1.2 Rumusan Masalah

1. Bagaimana validitas dan reliabilitas butir soal tes sumatif yang dibuat guru mata pelajaran IPA di SMP Xaverius 1 Jambi?
2. Bagaimana indeks kesulitan butir soal tes sumatif yang dibuat guru mata pelajaran IPA di SMP Xaverius 1 Jambi?
3. Bagaimana indeks daya pembeda butir soal tes sumatif yang dibuat guru mata pelajaran IPA di SMP Xaverius 1 Jambi?

1.3 Tujuan Penelitian

1. Menganalisis validitas dan reliabilitas butir soal tes sumatif yang dibuat guru mata pelajaran IPA di SMP Xaverius 1 Jambi?
2. Menganalisis indeks kesulitan butir soal tes sumatif yang dibuat guru mata pelajaran IPA di SMP Xaverius 1 Jambi?
3. Menganalisis indeks daya pembeda butir soal tes sumatif yang dibuat guru mata pelajaran IPA di SMP Xaverius 1 Jambi?

1.4 Manfaat Penelitian

Hasil pengujian analisis butir-butir soal sumatif yang diberikan kepada peserta didik di SMP Xaverius 1 Jambi akan menjadi bahan pertimbangan bagi Stakeholder dan guru-guru untuk melakukan evaluasi terhadap butir-butir soal yang dibuat.

Penelitian ini juga memberikan keterampilan yang sangat bermanfaat bagi calon guru, agar memiliki kemampuan untuk melakukan analisis butir-butir soal yang dikembangkan. Keterampilan ini tentu akan sangat berguna ketika sudah berada dilakukan profesional sebagai seorang guru, mengingat dalam proses pembelajaran tes formatif dan sumatif sangat penting.