

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Evaluasi merupakan salah satu tahap yang sangat penting dalam proses pembelajaran. Di dalam evaluasi, terdapat proses memahami, menilai, mendapatkan, dan mengomunikasikan suatu informasi untuk pengambilan keputusan. Evaluasi bertujuan untuk mengetahui taraf ketercapaian pembelajaran siswa yang disesuaikan dengan kurikulum. Selain terhadap siswa, evaluasi juga bertujuan untuk menentukan efektivitas kinerjanya guru dan lembaga pendidikan, serta sebagai pedoman untuk perbaikan sistem pembelajaran di masa depan. Hal inilah yang menyebabkan tahap evaluasi menjadi tahap yang sangat penting dalam dunia pendidikan.

Kualitas evaluasi hasil belajar siswa merupakan salah satu faktor yang menentukan kualitas pendidikan. Hal ini dijelaskan dalam Undang-undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional Bab XVI Pasal 57 ayat 1, bahwa evaluasi dilakukan dalam rangka pengendalian mutu pendidikan secara nasional sebagai akuntabilitas penyelenggaraan pendidikan kepada pihak-pihak terkait. Hasil akhir dari evaluasi haruslah memperoleh data dan informasi yang dapat dijadikan sebagai pedoman bagi penyelenggara pendidikan dalam pengambilan keputusan terhadap peserta didik (Sarea dan Hadi, 2015). Namun, pada kenyataannya masih banyak dijumpai hasil evaluasi dalam suatu lembaga pendidikan tidak menggambarkan dan membedakan kemampuan siswa secara tepat. Misalnya, sering dijumpai di lapangan bahwa hasil evaluasi siswa yang berkemampuan biasa-biasa saja, lebih tinggi daripada siswa yang biasanya dianggap guru berkemampuan lebih. Keputusan yang diambil terhadap kemampuan peserta didik untuk dicantumkan dalam laporan akhir sekolah, seringkali harus melalui berbagai pertimbangan subjektif dari guru terhadap peserta didik, sehingga evaluasi hasil belajarnya menjadi kurang konsisten. Konsistensi hasil evaluasi ini tidak terlepas dari kualitas proses yang melatarbelakanginya, yakni pengukuran.

Pengukuran identik dengan proses pemberian label berupa angka (kuantifikasi) terhadap siswa yang diharapkan dapat menunjukkan kemampuan siswa dalam setiap mata pelajaran (Sarea dan Hadi, 2015). Hal ini didasari pada penampakan siswa baik berupa motivasi, prestasi, maupun kepercayaan diri siswa yang kemudian diakumulasikan dalam bentuk angka. Dalam kegiatan ini diperlukan alat atau instrumen penilaian, agar pelaksanaannya lebih terarah. Dengan kata lain, kualitas proses pengukuran tersebut, sangat dipengaruhi oleh kualitas alat ukur yang digunakan. Dalam Rahmani (2015) dikatakan bahwa instrumen yang digunakan untuk pengumpulan data dapat berupa tes atau nontes.

Teknik penilaian tes dilakukan dengan cara memberikan seperangkat soal atau pertanyaan kepada peserta didik untuk mengukur aspek perilaku peserta didik. Penilaian seharusnya tidak sekedar memberi soal pada peserta didik kemudian selesai, tetapi guru harus menindaklanjutinya untuk kepentingan pembelajaran. Untuk itu, guru memerlukan instrumen berupa soal-soal yang baik untuk dapat mengukur hasil belajar peserta didik dengan baik dan akurat (Budiman, 2014 ; Sarea dan Hadi, 2015). Sebuah instrumen penilaian sebelum diujikan kepada peserta didik, harus terlebih dahulu diuji kualitasnya, karena berpengaruh langsung terhadap keakuratan status hasil pencapaian belajar peserta didik.

Salah satu bentuk penilaian hasil belajar peserta didik yang diselenggarakan di sekolah yaitu ujian tengah semester. Sejauh ini, ujian tengah semester di sekolah pada umumnya menggunakan soal berbentuk *multiple choice* (Arikunto, 2017). Tes ini terdiri atas satu pertanyaan dan beberapa pilihan jawaban. Siswa yang memiliki kemampuan rendah akan sulit untuk menebak karena pilihan jawaban terdiri dari jawaban benar dan jawaban pengecoh. Semakin banyak pilihan jawaban maka semakin banyak pengecoh sehingga kemungkinan besar siswa yang menebak akan salah (Khaerudin, 2016). Keberadaan pengecoh dalam soal *multiple choice* dapat dijadikan sebagai acuan untuk mengukur kemampuan siswa sekaligus alat ukur kualitas soal sebagai instrumen penilaian.

Kualitas soal yang baik harus memiliki validitas, reabilitas, daya pembeda, tingkat kesukaran dan pengecoh (*distractor*) yang baik (Purwanti, 2014; Wati, 2015). Irmalasari (2016) dalam penelitiannya juga mengatakan bahwa instrumen penilaian yang baik akan memotivasi guru menggunakan strategi mengajar yang

lebih menarik dan memotivasi siswa untuk belajar lebih giat. Dengan kata lain, untuk meningkatkan kualitas belajar siswa diperlukan peningkatan kualitas instrumen penilaian juga, misalnya dengan menyusun soal yang memenuhi syarat kualitas yang baik.

Beberapa ahli mengatakan, soal yang baik berarti soal yang tidak terlalu mudah atau tidak terlalu sukar. Menurut Arikunto (2017) soal yang terlalu mudah akan menyebabkan kemampuan siswa berpikir kritis sedikit menurun. Sedangkan, penelitian lain mengungkapkan soal yang terlalu sulit dapat menyebabkan siswa gampang menyerah dan tidak termotivasi untuk memecahkannya dengan baik. Selain itu, soal yang baik seharusnya juga dapat membedakan antara siswa yang telah memahami materi dengan siswa yang belum memahami materi. Namun, pada umumnya, soal-soal yang digunakan sekolah-sekolah dalam ujian, sering tidak memenuhi kriteria tersebut. Hal ini berkaitan erat dengan prosedur pembuatan soal tersebut. Setelah soal disusun, tidak dapat langsung diujikan kepada siswa, melainkan harus melalui tahap analisis butir soal untuk memastikan kualitasnya terlebih dahulu.

Nophitalia (2010) menemukan kasus dalam penelitiannya, dimana kualitas soal yang dibuat guru, sangat memprihatinkan. Pada pembuatan soal, guru lebih banyak memilih cara instan yaitu menyalin soal dari sumber buku teks atau lembar kerja sekolah, dari pada membuat soal sendiri. Kemudian soal itu langsung diujikan tanpa dilakukan analisis terlebih dahulu. Hal seperti ini juga ditemukan penulis dalam pengamatan di lapangan saat menjalani program latihan mengajar (PPL) di sekolah. Ketika hal seperti ini dilaksanakan, hasil evaluasi yang didapatkan oleh siswa sering menjadi masalah, misalnya rata-rata nilai ujian siswa rendah (tidak mencapai KKM).

Berdasarkan penelitian sebelumnya yang dilakukan oleh Suryawati dan Zulfikar (2012) mengenai analisis soal sebagai alat evaluasi hasil belajar, didapatkan hasil : (1) 28% soal pilihan ganda tergolong sukar dan 72% tergolong sedang. Hal ini berarti soal dapat dikatakan baik, (2) 32% daya beda tergolong jelek, 20% tergolong cukup, 16% tergolong baik, dan 28% tergolong sangat jelek dengan daya beda bertanda negatif. Artinya, soal tidak dapat membedakan siswa yang pandai dan kurang pandai. Hal ini juga ditemukan dalam penelitian Halik (2017)

mengenai analisis butir soal UAS, dimana kualitas soal pilihan ganda yang dianalisis menunjukkan kualitas : (1) Berdasarkan kriteria tingkat kesukaran soal diperoleh soal sangat mudah 2 butir (5%), sedang 24 butir (60%), sukar 12 butir (30%), dan sangat sukar 2 butir (5%), (2) Berdasarkan kriteria daya beda soal diperoleh soal dengan daya beda sangat jelek 3 butir (7,5%), jelek 12 butir (30%), cukup 14 butir (35%), dan baik 11 butir (27,5%). (3) Berdasarkan keefektifan pengecoh diperoleh 28 butir (70%) dalam kategori sangat baik, 9 butir (22,5%) dalam kategori baik, dan 3 butir (7,5%) dalam kategori buruk. Di akhir, disimpulkan bahwa terdapat 15 soal berkualitas baik dan 25 soal yang kurang baik.

Berdasarkan hasil wawancara dengan salah satu guru Biologi di SMA Negeri 1 Pematangsiantar, guru kesulitan mendeteksi kemampuan siswa yang sebenarnya berdasarkan hasil ujian karena didapati banyak pola jawaban siswa yang mirip. Padahal, ideal soal yang baik, harus dapat membedakan kemampuan siswa. Selain itu, distribusi nilai siswa juga kurang bervariasi. Saat proses ujian UTS berlangsung, keadaan siswa kurang kondusif, karena siswa sering menanyakan apa maksud beberapa soal dalam materi sistem pencernaan dan pernafasan. Belum dapat dipastikan, hal ini disebabkan oleh penggunaan bahasa yang sulit atau karena faktor lain. Telah dilakukan observasi awal ke sekolah, tentang penyusunan soal *multiple choice* pada tahun-tahun sebelumnya. Guru mata pelajaran Biologi menyatakan bahwa, proses analisis butir soal belum pernah dilakukan saat akan menyusun soal ujian. Guru juga belum terlalu memperhatikan distribusi jenis soal C1-C6 dalam menyusun soal, sehingga belum dapat dipastikan juga mengenai taraf kesukaran soal yang berpengaruh terhadap hasil pencapaian belajar siswa. Untuk itulah perlu diadakan analisis butir soal sebelum diujikan pada siswa, agar guru dapat membuat kesimpulan yang lebih akurat mengenai taraf kemampuan masing-masing peserta didik.

Berdasarkan uraian di atas, maka peneliti akan melakukan penelitian yang berjudul “Analisis Butir Soal *Multiple Choice* pada Ujian Tengah Semester Genap Kelas XI Mata Pelajaran Biologi di SMA Negeri 1 Pematangsiantar Tahun Pembelajaran 2018/2019”.

1.2. Identifikasi Masalah

Berdasarkan latar belakang masalah yang dikemukakan diatas, maka dapat diambil pokok-pokok masalah sebagai berikut :

1. Soal Ujian Tengah Semester Genap Mata Pelajaran Biologi Kelas XI IPA SMAN 1 Pematangsiantar Tahun Pembelajaran 2018/2019 belum dianalisis secara kuantitatif, dari segi tingkat kesukaran, daya pembeda, dan efektivitas pengecoh
2. Soal Ujian Tengah Semester Genap Mata Pelajaran Biologi Kelas XI IPA SMAN 1 Pematangsiantar Tahun Pembelajaran 2018/2019 belum dianalisis secara kualitatif, dari segi validitas isi (materi, konstruk dan bahasa)
3. Persebaran tipe soal berdasarkan ranah kognitif Taksonomi Bloom pada soal Ujian Tengah Semester Genap Mata Pelajaran Biologi Kelas XI IPA SMAN 1 Pematangsiantar Tahun Pembelajaran 2018/2019 belum diketahui secara pasti.

1.3. Batasan Masalah

Berdasarkan identifikasi masalah yang diuraikan, maka penulis membatasi masalah dalam penelitian ini. Batasan masalah dalam penelitian ini adalah:

1. Analisis kualitas butir soal Ujian Tengah Semester Genap Mata Pelajaran Biologi Kelas XI IPA SMAN 1 Pematangsiantar Tahun Pembelajaran 2018/2019 dari segi kuantitatif, meliputi aspek tingkat kesukaran, daya pembeda dan efektivitas pengecoh soal
2. Analisis kualitas butir soal Ujian Tengah Semester Genap Mata Pelajaran Biologi Kelas XI IPA SMAN 1 Pematangsiantar Tahun Pembelajaran 2018/2019 dari segi kualitatif, meliputi validitas isi dan persebaran tipe soal berdasarkan ranah kognitif Taksonomi Bloom.

1.4. Rumusan Masalah

Berdasarkan batasan masalah yang telah dipaparkan di atas, maka dapat dirumuskan bahwa masalah dalam penelitian ini adalah :

1. Bagaimana kualitas butir soal ujian tengah semester genap mata pelajaran Biologi kelas XI SMAN 1 Pematangsiantar tahun pembelajaran 2018/2019 dari segi tingkat kesukaran, daya pembeda, dan efektivitas pengecohnya ?

2. Bagaimana kualitas butir soal ujian tengah semester genap mata pelajaran Biologi kelas XI SMA N 1 Pematangsiantar tahun pembelajaran 2018/2019 dari segi validitas isi ?
3. Bagaimana persebaran tipe soal ujian tengah semester genap mata pelajaran Biologi kelas XI SMA N 1 Pematangsiantar tahun pembelajaran 2018/2019 berdasarkan ranah kognitif Taksonomi Bloom ?

1.5. Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan dalam penelitian ini adalah untuk mengetahui tentang :

1. Kualitas butir soal ujian tengah semester genap mata pelajaran Biologi kelas XI SMA N 1 Pematangsiantar tahun pembelajaran 2018/2019 dari segi tingkat kesukaran, daya pembeda dan efektivitas pengecoh.
2. Kualitas butir soal ujian tengah semester genap mata pelajaran Biologi kelas XI SMA N 1 Pematangsiantar tahun pembelajaran 2018/2019 dari segi validitas isi
3. Persebaran tipe soal ujian tengah semester genap mata pelajaran Biologi kelas XI SMA N 1 Pematangsiantar tahun pembelajaran 2018/2019 berdasarkan ranah kognitif Taksonomi Bloom .

1.6. Manfaat Penelitian

Berdasarkan dari tujuan penelitian di atas, maka penelitian ini diharapkan dapat memberi manfaat sebagai berikut :

1. Hasil penelitian menjadi sumber acuan untuk penelitian lanjutan dan konsep yang dihasilkan dapat menjadi masukan bagi dunia pendidikan, khususnya bidang evaluasi pembelajaran
2. Memberi masukan kepada lembaga pendidikan tentang cara menganalisis mutu instrumen evaluasi serta pertimbangan dalam meningkatkan kualitas evaluasi pembelajaran
3. Menambah pengetahuan dan pengalaman dalam menganalisis kualitas tes atau alat evaluasi pembelajaran serta mengetahui cara untuk memperbaikinya bagi peneliti dan masyarakat.