

**ANALISIS SOAL ULANGAN AKHIR SEMESTER (UAS) BIOLOGI KELAS X IPA SMA
NEGERI 5 KOTA TANJUNGPINANG
TAHUN AJARAN 2020/2021**

Irfan Saputra¹, Nevrita², Nurul Asikin³
saputrairfan209@gmail.com¹

Program studi Pendidikan Biologi, Fakultas Keguruan dan Ilmu Pendidikan,
Universitas Maritim Raja Ali Haji

ABSTRACT

This study aims to determine the quality of the items in the Class X Science Biology Final Semester of SMA Negeri 5 Tanjungpinang City for the 2020/2021 academic year in terms of validity, reliability, level of difficulty, discriminatory power and effectiveness of distractors. This research is quantitative descriptive. The subjects in this study were students of class X IPA SMA Negeri 5 Tanjungpinang City. The technique of data collection is done by the method of documentation. The data obtained were analyzed using the program Anates Version 4.0.9. The results of the study show: (1) The quality of the questions is based on validity, 32 questions or 80% are declared valid and 8 questions or 20% are declared invalid (2) The quality of the questions is based on reliability, the reliability coefficient is 0.86, these results shows that the questions have high reliability (3) The quality of the questions is based on the level of difficulty, there are 16 questions or 40% of very easy questions, 14 questions or 35% of easy questions, 4 items or 10% of moderate questions, 1 item or 2.5% difficult questions, and 5 items or 12.5% very difficult questions (4) The quality of the questions is based on discriminating power, there are 6 items or 15% poor discriminating power, 18 items or 45% sufficient discriminatory power, 14 items or 35% good discriminating power, 0 items or 0% excellent discriminating power, and 2 items or 5% discriminatory power Negative (not good) (5) Quality of questions based on effectiveness analysis pe fooling, obtained 3 items or 7.5% very good distractor effectiveness, 8 items or 20% good distractor effectiveness, 12 items or 30% sufficient distractor effectiveness, 12 items or 30% effectiveness bad distractors, 5 items or 12.5% the effectiveness of distractors that are not good or 25%, questions that are not good amount to 11 items or 27.5%, and questions with very poor quality are 2 items or 5 %.

Keywords : Analysis of Final Semester Exams, Biology, SMA Negeri 5 Tanjung Pinang City

I. Pendahuluan

Salah satu kompetensi yang harus dimiliki seorang pendidik adalah kemampuan melaksanakan evaluasi, baik dalam proses pembelajaran maupun penilaian hasil belajar. Kemampuan melaksanakan evaluasi pembelajaran merupakan kemampuan dasar yang mesti dikuasai oleh seorang pendidik maupun calon pendidik sebagai salah satu kompetensi profesionalnya.

Evaluasi pembelajaran merupakan satu kompetensi profesional seorang pendidik. Kompetensi tersebut sejalan dengan instrumen penilaian kemampuan guru, yang salah satu indikatornya adalah

melakukan evaluasi pembelajaran. Istilah evaluasi pembelajaran sering disamaartikan dengan ujian. Meskipun saling berkaitan, akan tetapi tidak mencakup keseluruhan makna yang sebenarnya. Secara umum tujuan evaluasi pembelajaran adalah mengetahui keefektifan dan efisiensi sistem pembelajaran secara luas. Hal ini sejalan dengan yang dikatakan Ratnawulan dan Rusdiana (2014 : 1) “ berhasil atau tidaknya pendidikan dalam mencapai tujuannya dapat dilihat setelah dilakukan evaluasi terhadap output atau lulusan yang dihasilkannya. Jika output lulusan, hasilnya sesuai dengan apa yang telah digariskan dalam tujuan pendidikan, maka usaha pendidikan itu dapat dinilai berhasil, tetapi jika sebaliknya, maka ia dinilai gagal.” Zainal Arifin (2014: 5) memaparkan bahwa “evaluasi adalah suatu proses yang sistematis dan berkelanjutan untuk menentukan kualitas (nilai dan arti) dari sesuatu, berdasarkan pertimbangan dan kriteria tertentu dalam rangka pembuatan keputusan”.

Dalam kegiatan evaluasi diperlukan alat penilaian baik berupa tes atau pun non tes. Menurut Asrul, dkk (2014: 48) “Instrumen evaluasi pembelajaran jenis tes adalah teknik yang paling umum digunakan dalam kegiatan pengukuran. Meskipun teknik ini tidak selalu terbaik dan tepat untuk beberapa tujuan”

Bentuk instrumen penilaian sebagai alat ukur hasil belajar siswa harus dapat dipertanggungjawabkan, artinya bahwa tes yang digunakan tersebut dapat memenuhi syarat sebagai alat evaluasi yang baik bila dilihat dari kualitas butir soal.

Ulangan Akhir Semester merupakan salah satu bentuk evaluasi untuk mengukur keberhasilan siswa/siswi dalam proses pembelajaran yang telah berlangsung selama satu semester. Sehingga seorang guru akan mempunyai acuan untuk menentukan materi-materi yang lain dengan mempertimbangkan strategi-strategi yang akan dilakukan pada saat pembelajaran berikutnya. Penting pula bagi seorang guru untuk membuat soal soal ulangan akhir semester yang berkualitas sehingga alat ukur tersebut dapat menjadi acuan yang baik. Salah satu cara untuk mengetahui kualitas soal ulangan yang guru berikan kepada peserta didik adalah dengan melakukan analisis butir soal.

Dari hasil wawancara penulis dengan guru mata pelajaran Biologi SMA Negeri 5 Tanjungpinang, bahwasannya masih banyak siswa yang mendapatkan nilai dibawah Kriteria Ketuntasan Minimal (KKM) yaitu dengan nilai ketuntasan sebesar 71, dari 71 orang siswa yang mengikuti ujian hanya 35 orang yang mencapai nilai ketuntasan tersebut, hal ini dapat dipengaruhi oleh kualitas soal ulangan yang diberikan oleh guru. Tujuan penelitian ini adalah untuk mengetahui kualitas butir soal ulangan Akhir Semester Biologi kelas X IPA SMA Negeri 5 Kota Tanjungpinang ditinjau dari segi validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektifitas pengecoh.

II. Metode Penelitian

Penelitian ini dilakukan di SMA Negeri 5 Kota Tanjungpinang pada kelas X IPA dengan jumlah populasi 71 siswa. Teknik pengambilan sampel pada penelitian ini menggunakan teknik sampling jenuh (sesnsus). Maka jumlah sampel yang digunakan adalah seluruh populasi yaitu sebanyak 71 siswa. Jenis penelitian ini merupakan penelitian deskriptif kuantitatif dimana data yang diperoleh berupa angka-angka atau pernyataan yang dinilai dan dianalisis secara statistik.

Rancangan penelitian terdiri dari tahap persiapan dengan melakukan observasi dilokasi penelitian dan berkonsultasi dengan guru mata pelajaran biologi, tahap pelaksanaan mengumpulkan data penelitian berupa soal dan jawaban ulangan siswa kemudian dianalisis menggunakan program Anates Versi 4.0.9, tahap akhir yaitu menuangkan hasil pengolahan data dan kesimpulan dari analisis data soal ulangan kedalam bentuk tulisan secara sistematis.

III. Hasil dan Pembahasan

Penelitian ini dilaksanakan dengan tujuan untuk mengetahui kualitas soal Ulangan Akhir Semester (UAS) biologi siswa kelas x SMA Negeri 5 Kota Tanjungpinang Tahun Ajaran 2020/2021. Kualitas soal tersebut ditinjau dari segi validitas, realibilitas, daya pembeda soal, tingkat kesukaran, dan efektifitas pengecoh. Soal Ulangan Akhir Semester yang dianalisis adalah bentuk soal pilihan ganda dengan jumlah 40 butir soal dengan peserta yang mengikuti ujian sebanyak 71 siswa.

Validitas merupakan suatu ukuran atau standar yang menunjukkan tingkat ketepatan atau kesahihan suatu instrumen. Menurut Asrul, dkk (2014 :121), “Valid artinya sah atau tepat, jadi tes yang valid berarti tes tersebut merupakan alat ukur yang tepat untuk mengukur suatu objek. Berdasarkan pengertian ini maka validitas tes pada dasarnya berkaitan dengan objek ketepatan dan kesesuaian antara tes sebagai alat ukur dengan objek yang diukur.” Berdasarkan 40 butir soal yang dianalisis didapatkan 32 soal atau sebesar 80% dinyatakan valid dan 8 soal atau sebesar 20% dinyatakan tidak valid. Dari hasil tersebut menunjukkan kualitas soal yang baik ditinjau dari segi validitasnya. Oleh karena itu, butir soal yang valid dapat disimpan di bank soal dan butir soal yang tidak valid perlu dilakukan revisi.

Realibilitas suatu soal dipengaruhi oleh jumlah butir soal yang digunakan, semakin banyak butir soal maka semakin tinggi realibilitasnya. Hal ini sejalan dengan teori Suharsimi Arikunto (2013: 101) yang menyatakan bahwa “tes yang terdiri dari banyak butir akan lebih valid daripada tes yang terdiri dari beberapa butir saja. Tinggi rendahnya tingkat validitas dapat menunjukkan tinggi rendahnya koefisien reliabilitas, sehingga semakin panjang tes maka reliabilitasnya semakin tinggi”. Berdasarkan hasil analisis, soal ujian yang dianalisis memiliki realibilitas sebesar 0,86, hasil ini menunjukkan soal tersebut memiliki reliabilitas yang tinggi karena nilai reliabilitas $\geq 0,70$.

Kualitas soal dapat dilihat dari tingkat kesukaran butir soal-soal tersebut, tingkat kesukaran adalah banyaknya siswa yang menjawab suatu soal dengan benar dengan jumlah seluruh peserta tes. Menurut Asrul, dkk (2014 : 149) “Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi, karena diluar jangkauannya”. Zulaiha (2008 : 14) menambahkan “makin besar tingkat kesukaran makin mudah soal tersebut begitu pula sebaliknya makin kecil tingkat kesukaran makin sukar soal tersebut.” Dari hasil perhitungan tingkat kesukaran didapatkan 16 butir soal atau sebesar 40% termasuk kriteria soal sangat mudah, 14 butir soal atau sebesar 35% kriteria soal mudah, 4 butir soal atau sebesar 10% kriteria soal sedang, 1 butir soal atau sebesar 2,5% kriteria soal sukar, dan 5 butir soal atau sebesar 12,5% tergolong kedalam kriteria soal sangat sukar. Berdasarkan hasil tersebut dapat disimpulkan soal tersebut memiliki tingkat kesukaran yang kurang baik, Hal ini dapat dilihat dari jumlah kriteria tingkat kesukaran soal sangat mudah terlalu mendominasi dibandingkan kriteria yang lainnya yaitu sebesar 40% soal. Berkaitan dengan hal tersebut agar tingkat kesukaran soal proposional perlu dilakukan perbaikan terhadap soal yang terlalu mudah agar dapat digunakan lagi untuk tes berikutnya.

Daya pembeda soal adalah kemampuan sesuatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi Asrul, dkk (2014 : 151). Hasil perhitungan daya pembeda didapatkan 6 butir soal atau sebesar 15% dengan kriteria daya pembeda jelek, 18 butir soal atau sebesar 45% termasuk kriteria daya pembeda cukup, 14 butir soal atau sebesar 35% termasuk kedalam kriteria daya pembeda baik, 0 butir soal atau sebesar 0% daya pembeda baik sekali, dan 2 butir soal atau sebesar 5% kriteria daya pembeda Negatif (tidak baik). Berdasarkan uraian hasil analisis daya pembeda soal dapat disimpulkan bahwa Soal memiliki daya pembeda

yang cukup baik, hal ini dapat dilihat dari sebagian besar daya pembeda yakni sebesar 45% termasuk kedalam kriteria daya pembeda cukup. Kemudian untuk butir soal yang memiliki daya pembeda baik yakni sebesar 35% dapat disimpan kedalam bank soal untuk digunakan pada tes berikutnya. Butir soal dengan daya pembeda jelek sebesar 15% perlu dilakukan perbaikan, dan untuk butir soal dengan kriteria negatif yakni sebesar 5% sebaiknya dibuang atau tidak digunakan lagi pada tes selanjutnya.

Efektifitas pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut telah dipilih paling sedikit 5% dari seluruh peserta tes. Dalam penelitian yang penulis lakukan jumlah peserta tes sebanyak 71 siswa, jadi 5% dari 71 adalah 3,55. Maka efektifitas pengecoh dapat dikatakan bekerja dengan baik apabila dipilih lebih dari 3 peserta tes. Menurut Aprianto (2008) terdapat beberapa faktor yang mempengaruhi berfungsi tidaknya pengecoh yaitu jika soal terlalu mudah, pokok soal memberi petunjuk pada kunci jawaban dan siswa sudah mengetahui materi yang akan ditanyakan terlalu mudah. Kemudian Widodo (2010) menambahkan “penyebab pengecoh yang tidak dipilih oleh peserta tes karena terlalu kelihatan menyesatkan.” Perhitungan efektifitas pengecoh dilakukan dengan menggunakan program Anates 4.0.9. Hasil perhitungan efektifitas pengecoh dari Anates 4.0.9 kemudian diinterpretasikan kedalam 5 kriteria yang diadaptasi dari skala *Likert*. Hasil analisis daya pembeda soal didapatkan 3 butir soal atau sebesar 7,5% memiliki efektifitas pengecoh yang sangat baik, 8 butir soal atau sebesar 20% memiliki efektifitas pengecoh yang baik, 12 butir soal atau sebesar 30% memiliki efektifitas pengecoh yang cukup, 12 butir soal atau sebesar 30% memiliki efektifitas pengecoh yang kurang baik, 5 butir soal atau sebesar 12,5% memiliki efektifitas pengecoh yang tidak baik. Berdasarkan hasil analisis tersebut, menunjukkan bahwa dilihat dari efektifitas pengecoh jawaban dapat dikategorikan cukup karena mendapatkan persentase sebesar 30% dari jumlah seluruh soal, untuk itu perlu dilakukan perbaikan. Butir soal dengan fungsi pengecoh sangat baik dan baik dapat dimasukkan kedalam bank soal untuk dapat digunakan pada tes berikutnya. Kemudian butir soal dengan fungsi pengecoh kurang baik dan tidak baik harus dilakukan perbaikan, cara yang dapat digunakan antara lain dengan mengganti pengecoh yang tidak berfungsi dengan pengecoh yang lebih mendekati atau lebih setara dengan kunci jawaban agar peserta tes dapat lebih berpikir dan mengasah kemampuannya.

Analisis butir soal adalah suatu prosedur yang sistematis, yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang kita susun. Menurut Nana Sudjana (2012: 135) “analisis butir soal adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai.” Perhitungan hasil analisis Validitas, Reliabilitas, Tingkat Kesukaran, Daya Pembeda, dan Sebaran Pengecoh menggunakan Program Anates 4.0.9, selanjutnya diinterpretasikan kualitas butir soal tersebut menggunakan skala *Likert* menjadi 5 kriteria yaitu sangat baik, baik, sedang, tidak baik, dan sangat tidak baik. Hasil analisis Kualitas Butir Soal didapatkan 8 butir soal atau sebesar 20% memiliki kriteria kualitas soal sangat baik, 6 butir soal atau sebesar 15% memiliki kriteria kualitas soal yang baik, 4 butir soal atau sebesar 10% memiliki kriteria kualitas soal sedang, 18 butir soal atau sebesar 45% memiliki kriteria kualitas soal tidak baik, dan 4 butir soal atau sebesar 10% memiliki kriteria kualitas soal sangat tidak baik. Tindak lanjut yang dapat dilakukan antara lain, butir soal yang memiliki kriteria sangat baik dapat dimasukkan kedalam bank soal untuk keperluan instrumen tes berikutnya. Butir soal dengan kriteria baik dan sedang dapat diperbaiki sesuai dengan indikator yang belum terpenuhi. Kemudian butir soal dengan kriteria tidak baik dan sangat tidak baik sebaiknya dibuang atau tidak digunakan lagi sebagai instrumen tes.

IV. Kesimpulan

Kualitas soal memiliki hasil yang baik ditinjau dari segi validitasnya, butir soal yang valid dapat disimpan di bank soal dan butir soal yang tidak valid perlu dilakukan revisi. Dari segi reliabilitas diketahui bahwa soal ujian yang dianalisis memiliki reliabilitas yang tinggi. Soal yang memiliki

reliabilitas tinggi menunjukkan apabila soal tersebut digunakan berulang-ulang kepada siswa maka hasil yang didapatkan tidak akan jauh berbeda. Dari hasil pengukuran tingkat kesukaran dapat disimpulkan soal memiliki tingkat kesukaran yang cukup karena didominasi oleh butir soal yang terlalu mudah sehingga perlu dilakukan revisi jika ingin digunakan lagi pada tes selanjutnya. Daya pembeda soal menunjukkan hasil yang cukup baik, dengan demikian soal tersebut dapat dikatakan cukup baik untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Berdasarkan efektifitas pengecoh, soal memiliki efektifitas pengecoh yang cukup, hasil ini menunjukkan pengecoh soal belum berfungsi dengan baik sehingga perlu dilakukan perbaikan dengan cara mengganti pengecoh dengan pengecoh yang lebih setara antara pilihan jawaban keseluruhan

V. Daftar Pustaka

- Aprianto. 2008. Kualitas Butir Tes Soal Pilihan Ganda. Banjarmasin. On line at <http://apri76.wordpress.com/2008/12/31/kualitas-butir-tes-soal-pilihan-ganda> (diakses tanggal 25 Mei 2021).
- Arikunto, S. 2013. Dasar-dasar Evaluasi Pendidikan. Jakarta: Bumi Askara.
- Asrul, Dkk. 2014. Evaluasi Pembelajaran. Medan: Citapustaka Media
- Elis Ratna Wulan dan H. A. Rusdiana. 2014. Evaluasi Pembelajaran. Yogyakarta: Multi Pressindo
- Nana Sudjana. 2012. Penilaian Hasil Proses Belajar Mengajar. Bandung: PT Remaja Rosdakarya.
- Widodo. 2010. Analisis Butir Soal Tes. Jurnal Pendidikan Penabur no 14/Tahun ke-9/Juni 2010.
- Zainal Arifin. 2014. Evaluasi Pembelajaran. Bandung: PT Remaja Rosdakarya.
- Zulaiha, Rahmah. 2008. Analisis Soal Secara Manual. Jakarta: Puspendik

VI. Ucapan Terimakasih

Peneliti mengucapkan terima kasih atas dukungan dari berbagai pihak yang membantu dalam penelitian ini, khususnya Dr. Hj. Nevrita, M.Pd., M.Si., dan Nurul Asikin, S.Pd., M.Pd., yang telah membantu selama penelitian dan proses penulisan artikel berlangsung.