

SKRIPSI

**ANALISIS KUALITAS BUTIR SOAL UAS (UJIAN AKHIR SEMESTER)
KELAS XI PADA MATA PELAJARAN KIMIA
DI SMA NEGERI 1 BENAI**

Diajukan Kepada Fakultas Ilmu Pendidikan dan Sains Islam Universitas Islam Kuantan Singingi
Untuk Memenuhi Salah Satu Syarat Guna Memperoleh Gelar Sarjana Pendidikan



OLEH

MAYA FEBRIANI CHANDRA
210309002

**PROGRAM STUDI PENDIDIKAN KIMIA
FAKULTAS ILMU PENDIDIKAN DAN SAINS ISLAM
UNIVERSITAS ISLAM KUANTAN SINGINGI
TELUK KUANTAN
1447 H/2025 M**

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Maya Febriani Chandra
Tempat/Tanggal Lahir : Simpang Tanah Lapang/ 22 Februari 2003
NPM : 210309002
Alamat : Simpang Tanah Lapang, Baserah
Program Studi : Pendidikan Kimia
Fakultas : Ilmu Pendidikan dan Sains Islam Universitas
Islam Kuantan Singingi

Dengan ini menyatakan bahwa skripsi yang saya buat dengan judul **“Analisis Kualitas Butir Soal UAS (Ujian Akhir Semester) Kelas XI Pada Mata Pelajaran Kimia Di SMA Negeri 1 Benai”** adalah benar karya saya sendiri dan bertanggung jawab atas data dan informasi yang termuat di dalamnya.

Demikian pernyataan ini saya buat dengan sesungguhnya, apabila dikemudian hari pernyataan saya terbukti tidak benar, maka saya bersedia menanggung semua resikonya.

Teluk Kuantan, 14 Juli 2025
Saya yang menyatakan


Maya Febriani Chandra
NPM. 210309002

DWI PUTRI MUSDANSI, S.Pd,M.Pd

DOSEN PROGRAM STUDI PENDIDIKAN KIMIA

FAKULTAS ILMU PENDIDIKAN DAN SAINS ISLAM

UNIVERSITAS ISLAM KUANTAN SINGINGI

NOTA DINAS

Perihal : Skripsi Maya Febriani Chandra

Kepada Yth,
Dekan Fakultas Ilmu Pendidikan dan
Sains Islam
Universitas Islam Kuantan Singingi
Di-

Teluk Kuantan

Assalamu'alaikum Warahmatullahi Wabarakatuh

Setelah membaca, meneliti, mengoreksi dan melakukan perbaikan terhadap skripsi saudara :

Nama : Maya Febriani Chandra
NPM : 210309002
Program Studi : Pendidikan Kimia
Fakultas : Fakultas Ilmu Pendidikan dan Sains Islam
Judul : Analisis Kualitas Butir Soal UAS (Ujian Akhir Semester)
Kelas XI Pada Mata Pelajaran Kimia Di SMA Negeri 1
Benai

Maka dengan ini dapat disetujui untuk diuji dan diberikan penilaian dalam sidang munaqasyah Program Studi Pendidikan Kimia, Fakultas Ilmu Pendidikan dan Sains Islam, Universitas Islam Kuantan Singingi.

Wasalamu'alaikum Warahmatullahi Wabarakatuh

Teluk Kuantan, 14 Juli 2025
Pembimbing I



Dwi Putri Musdansi, S.Pd,M.Pd
NIDN.1019049801

JUMRIANA RAHAYU NINGSIH, S.Pd, M.Si
DOSEN PROGRAM STUDI PENDIDIKAN KIMIA
FAKULTAS ILMU PENDIDIKAN DAN SAINS ISLAM
UNIVERSITAS ISLAM KUANTAN SINGINGI

NOTA DINAS

Perihal : Skripsi Maya Febriani Chandra

Kepada Yth,
Dekan Fakultas Ilmu Pendidikan dan
Sains Islam
Universitas Islam Kuantan Singingi
Di-

Teluk Kuantan

Assalamu'alaikum Warahmatullahi Wabarakatuh

Setelah membaca, meneliti, mengoreksi dan melakukan perbaikan terhadap skripsi saudara :

Nama : Maya Febriani Chandra
NPM : 210309002
Program Studi : Pendidikan Kimia
Fakultas : Fakultas Ilmu Pendidikan dan Sains Islam
Judul : Analisis Kualitas Butir Soal UAS (Ujian Akhir Semester)
Kelas XI Pada Mata Pelajaran Kimia Di SMA Negeri 1
Benai

Maka dengan ini dapat disetujui untuk diuji dan diberikan penilaian dalam sidang munaqasyah Program Studi Pendidikan Kimia, Fakultas Ilmu Pendidikan dan Sains Islam, Universitas Islam Kuantan Singingi.

Wasalamu'alaikum Warahmatullahi Wabarakatuh

Teluk Kuantan, 14 Juli 2025
Pembimbing II



Jumriana Rahayu Ningsih, S.Pd, M.Si
NIDN. 1013077803

PERSETUJUAN PEMBIMBING DAN KETUA PRODI

Skripsi dengan judul "Analisis Kualitas Butir Soal UAS (Ujian Akhir Semester) Kelas XI Pada Mata Pelajaran Kimia Di SMA Negeri 1 Benal" yang ditulis oleh Maya Febriani Chandra, NPM. 210309002 dapat diterima dan telah disetujui untuk diajukan dalam sidang munaqasyah Sarjana Strata Satu (S1) Program Studi Pendidikan Kimia Fakultas Ilmu Pendidikan dan Sains Islam Universitas Islam Kuantan Singingi untuk memenuhi salah satu syarat meraih gelar Sarjana Pendidikan (S.Pd).

Teluk Kuantan, 14 Juli 2025

Menyetujui

Pembimbing I



Dwi Putri Musdansi, S.Pd,M.Pd
NIDN.1019049801

Pembimbing II



Jumriana Rahayu Ningsih, S.Pd,M.Si
NIDN. 1013077803

Mengetahui,
Ketua Prodi Pendidikan Kimia


Rosa Murwindra, S.Pd., M.Si
NIDN. 1014078503

PENGESAHAN PENGUJI

Skripsi dengan judul “Analisis Kualitas Butir Soal UAS (Ujian Akhir Semester) Kelas XI Pada Mata Pelajaran Kimia Di SMA Negeri 1 Benai” yang ditulis oleh Maya Febriani Chandra, NPM. 210309002 telah diuji dalam sidang munaqasyah Program Studi Pendidikan Kimia Fakultas Ilmu Pendidikan dan Sains Islam Universitas Islam Kuantan Singingi. Skripsi ini telah diterima sebagai syarat untuk memperoleh gelar Sarjana Pendidikan (S.Pd) pada Program Studi Pendidikan Kimia.

Teluk Kuantan, 14 Juli 2025

Mengesahkan
Tim Sidang Munaqasyah
Ketua



Bustanur, S.Ag., M.Us
NIDN. 2120067501

Moderator



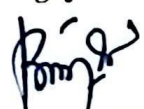
Dwi Putri Musdansi, S.Pd., M.Pd
NIDN. 1019049801

Secretaris



Jumriana Rahayu N., S.Pd., M.Si
NIDN. 1013077803

Penguji I



Rosa Murwindra, S.Pd., M.Si
NIDN. 1014078503

Penguji II



Dr. Irfandi, S.Pd., M.Pd
NIDN. 1012059601

Dekan Fakultas Ilmu Pendidikan dan Sains Islam
Universitas Islam Kuantan Singingi



Bustanur, S.Ag., M.Us
NIDN. 2120067501

v

ABSTRAK

MAYA FEBRIANI CHANDRA (2025) : “Analisis Kualitas Butir Soal UAS (Ujian Akhir Semester) Kelas XI Pada Mata Pelajaran Kimia Di SMA Negeri 1 Benai”

Penelitian ini bertujuan untuk menganalisis kualitas butir soal Ujian Akhir Semester (UAS) mata pelajaran Kimia kelas XI di SMA Negeri 1 Benai Tahun Ajaran 2023/2024 berdasarkan indikator validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh. Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan data sekunder berupa hasil ujian dari 106 siswa yang dianalisis menggunakan program ITEMAN. Hasil analisis menunjukkan bahwa 76% soal memiliki validitas tinggi hingga sangat tinggi, dengan reliabilitas *Alpha Cronbach* sebesar 0,793 yang tergolong tinggi. Daya pembeda soal menunjukkan 52% berada dalam kategori baik, dan tingkat kesukaran didominasi oleh soal kategori sedang (52%), sukar (40%), dan mudah (8%). Sebanyak 64% soal memiliki pengecoh yang efektif, sementara sisanya perlu revisi. Temuan ini menunjukkan bahwa sebagian besar soal telah memenuhi standar kualitas evaluasi, meskipun beberapa butir tetap perlu diperbaiki untuk meningkatkan akurasi dan objektivitas pengukuran.

Kata Kunci: daya pembeda, pengecoh, reliabilitas, tingkat kesukaran, validitas

ABSTRACT

MAYA FEBRIANI CHANDRA (2025) : "An Analysis of the Quality of Final Semester Examination Items for Grade XI in Chemistry Subject at SMA Negeri 1 Benai"

This study aims to analyze the quality of final semester examination (UAS) test items in Chemistry for eleventh-grade students at SMA Negeri 1 Benai in the 2023/2024 academic year based on indicators of validity, reliability, discrimination index, difficulty level, and distractor effectiveness. Using a descriptive quantitative approach, the study analyzed secondary data from 106 students' test results through the ITEMAN program. The results show that 76% of the items have high to very high validity, with a Cronbach's Alpha reliability of 0.793, indicating high consistency. The discrimination index shows that 52% of the items are in the good category, while item difficulty is dominated by moderate (52%), difficult (40%), and easy (8%) questions. Moreover, 64% of the items have effective distractors, while the rest require revision. These findings indicate that most test items meet quality standards for evaluation, although some require improvement to enhance the accuracy and objectivity of student assessment.

Keywords: discrimination index, distractor, reliability, difficulty level, validity

DAFTAR ISI

	Halaman
SURAT PERNYATAAN.....	i
NOTA DINAS PEMBIMBING	ii
LEMBAR PERSETUJUAN PEMBIMBING DAN KETUA PRODI	iv
LEMBAR PENGESAHAN PENGUJI	v
MOTTO	vi
PERSEMBAHAN.....	vii
ABSTRAK	viii
KATA PENGANTAR	x
DAFTAR ISI.....	xiii
DAFTAR TABEL.....	xv
DAFTAR GAMBAR.....	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	
A. Latar Belakang Masalah.....	1
B. Identifikasi Masalah	4
C. Batasan Masalah	5
D. Rumusan Masalah	5
E. Tujuan Penelitian	5
F. Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA	
A. Kajian Teoritis.....	8

B. Penelitian Relevan.....	17
C. Kerangka Konseptual	20

BAB III METODOLOGI PENELITIAN

A. Jenis Penelitian	23
B. Waktu dan lokasi penelitian	23
C. Populasi dan Sampel	24
D. Teknik Pengumpulan Data	24
E. Teknik Analisa Data.....	25

BAB IV PENYAJIAN DAN ANALISIS DATA

A. Gambaran Umum Lokasi Penelitian.....	35
B. Hasil Penelitian	37
C. Pembahasan	47

BAB V KESIMPULAN DAN SARAN

A. Kesimpulan	59
B. Saran	59

DAFTAR KEPUSTAKAAN

LAMPIRAN

DAFTAR TABEL

	Halaman
Tabel 2.1 Tabel Penelitian Relevan.....	20
Tabel 4.1 Identitas Sekolah Lokasi Penelitian.....	35
Tabel 4.2 Visi dan Misi Sekolah	37
Tabel 4.3 Reliabilitas Butir Soal.....	39

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Bagan Kerangka Konseptual	22
Gambar 4.1 Presentase Validitas Soal.....	38
Gambar 4.3 Presentase Daya Beda Soal.....	40
Gambar 4.4 Presentase Tingkat Kesukaran Soal.....	41
Gambar 4.5 Presentase Efektivitas Pengecoh.....	42

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Transkrip Wawancara	65
Lampiran 2 Dokumentasi Wawancara	67
Lampiran 3 Soal UAS Kimia Kelas XI	68
Lampiran 4 Silabus Kimia Kelas XI Semester 2.....	71
Lampiran 5 Kunci Jawaban Soal UAS Kimia Kelas XI.....	78
Lampiran 6 Lembar Jawaban Siswa Kelas XI.....	79
Lampiran 7 Surat Balasan Izin Riset	80
Lampiran 8 Rekapitulasi Jawaban Siswa Kelas XI	81
Lampiran 9 Output Iteman	86
Lampiran 10 Rekapitulasi Validitas Butir Soal	92
Lampiran 11 Rekapitulasi Daya Pembeda Soal	93
Lampiran 12 Rekapitulasi Tingkat Kesukaran Soal	94
Lampiran 13 Rekapitulasi Efektifitas Pengecoh	95
Lampiran 14 Rekapitulasi Kelayakan Soal.....	96

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Kualitas pendidikan menjadi perhatian utama setiap negara karena berkontribusi pada peningkatan kualitas sumber daya manusia, yang diharapkan dapat meningkatkan kesejahteraan masyarakat¹. Kualitas pendidikan sering diukur melalui kemampuan lulusan dalam aspek pengetahuan, keterampilan, dan perilaku, yang dicapai melalui proses pembelajaran di sekolah. Pembelajaran hakikatnya adalah usaha sadar dari seorang guru untuk membelajarkan siswanya (mengarahkan interaksi siswa dengan sumber belajar lainnya) dalam rangka mencapai tujuan yang diharapkan.² Salah satu cara untuk mengevaluasi keberhasilan pembelajaran adalah melalui evaluasi sumatif, seperti Ujian Akhir Semester (UAS). Evaluasi ini memerlukan keterampilan guru yang profesional agar dapat mendidik peserta didik menjadi individu yang unggul dan berkualitas.

Ujian akhir semester merupakan bagian penting dalam evaluasi pembelajaran di tingkat SMA, karena berfungsi untuk mengukur sejauh mana siswa memahami materi yang diajarkan. Dalam konteks pembelajaran kimia, evaluasi ini memiliki peran krusial dalam menilai

¹ T Anggraeni, —Analisis tingkat kesukaran soal ujian akhir pada mata pelajaran IPA,|| *Jurnal Penelitian dan Evaluasi Pendidikan* 7, no. 3 (2019): 197–205.

² S Sudarsono, —Analisis kualitas prediksi soal ujian nasional hasil diklat teknis substantif UN guru matematika MTs ditinjau dari tingkat kesukaran, daya pembeda dan kualitas pengecoh,|| *Widyadewata* 4, no. 1 (2022): 47–54.

penguasaan konsep, keterampilan analisis, serta kemampuan siswa dalam menerapkan teori kimia dalam berbagai situasi.³ Agar hasil evaluasi ini mencerminkan kemampuan siswa yang sebenarnya, kualitas soal ujian harus memenuhi standar tertentu, seperti keseimbangan tingkat kesulitan, daya pembeda yang baik, dan pengecoh yang efektif. Dalam pembelajaran kimia, penting untuk memastikan bahwa soal tidak hanya menguji hafalan tetapi juga pemahaman konsep, keterampilan berpikir kritis, dan pemecahan masalah. Faktor-faktor hasil belajar, seperti motivasi, minat, dan lingkungan belajar, sangat dipengaruhi oleh kualitas soal ujian ini.⁴ Misalnya, soal yang terlalu sulit atau tidak sesuai dengan kompetensi siswa dapat menurunkan motivasi dan rasa percaya diri mereka. Sebaliknya, soal yang dirancang dengan baik dan sesuai dengan tingkat pemahaman siswa dapat meningkatkan minat belajar serta menciptakan evaluasi yang lebih akurat, sehingga mendukung peningkatan kualitas pembelajaran kimia di sekolah.

Hasil wawancara dengan guru mata pelajaran Kimia di SMA Negeri 1 Benai menguatkan pentingnya kualitas soal dalam evaluasi pembelajaran. Guru mengungkapkan adanya ketidaksesuaian antara ekspektasi terhadap siswa dan hasil belajar yang diperoleh, di mana beberapa siswa yang biasanya menunjukkan performa baik selama proses pembelajaran justru mendapatkan nilai rendah pada ujian. Hal ini menunjukkan bahwa kualitas soal, metode belajar siswa, serta tekanan saat

³ A Wibowo, —Analisis validitas butir soal ujian sekolah pada mata pelajaran kimia, *Jurnal Pendidikan Sains* 7, no. 1 (2019): 25–35.

⁴ *Ibid.*

ujian menjadi faktor penting yang memengaruhi hasil evaluasi. Dengan demikian, kualitas soal ujian yang baik tidak hanya penting untuk memastikan akurasi evaluasi, tetapi juga untuk mendukung keberhasilan belajar siswa secara keseluruhan.

Kualitas soal ujian memiliki peran yang sangat penting dalam mengukur kemampuan akademik siswa secara objektif. Dalam penyusunan soal ujian objektif untuk mata pelajaran Kimia, diperlukan analisis yang mendalam untuk memastikan bahwa soal-soal tersebut memenuhi standar kualitas yang ideal. Kualitas soal yang baik ditentukan oleh beberapa faktor utama, seperti tingkat kesukaran soal, daya pembeda, efektivitas pengecoh (*distractor*), dan reliabilitas.⁵ Butir soal harus memiliki standar kualitas yang tinggi agar hasil ujian dapat mencerminkan kemampuan siswa dengan akurat serta memberikan data yang valid bagi guru untuk mengevaluasi proses pembelajaran.

Penelitian oleh Sridadi, dkk. yang berjudul: —*Evaluasi tes hasil belajar ulangan akhir semester genap mata pelajaran PJOK kelas VIII SMP N 1 Ngemplak tahun ajaran 2017/2018 dengan analisis butir soal* —, menunjukkan bahwa soal dengan tingkat kesukaran sedang dan daya pembeda yang baik memberikan penilaian yang lebih akurat terhadap kemampuan siswa.⁶ Sementara itu, penelitian yang dilakukan oleh Amelia

⁵ A Wibowo, —Analisis validitas butir soal ujian sekolah pada mata pelajaran kimia,|| *Jurnal Pendidikan Sains* 7, no. 1 (2019): 25–35.

⁶ Sridadi, Riky Dwihandaka, dan Ariyo Bagiastomo, —Evaluasi tes hasil belajar ulangan akhir semester genap mata pelajaran PJOK kelas VIII SMP N 1 Ngemplak tahun ajaran 2017/2018 dengan analisis butir soal,|| *Jurnal Pendidikan Jasmani Indonesia* 16, no. 1 (30 April 2020): 28–40, <https://doi.org/10.21831/jpji.v16i1.31416>.

et al. yang berjudul: —*Chemistry Learning Outcomes Assessment: How is The Quality of The Tests Made by The Teacher?*|| menyebutkan bahwa soal yang dibuat oleh guru di SMA menunjukkan kualitas yang cukup baik, dengan tingkat kesukaran yang seimbang dan reliabilitas yang tinggi. Soal ideal memiliki tingkat kesulitan sedang, daya pembeda yang baik, dan pengecoh yang efektif untuk mencegah jawaban asal tebak.⁷

Berdasarkan latar belakang inilah maka penulis bermaksud melakukan penelitian tentang —Analisis Kualitas Butir Soal UAS (Ujian Akhir Semester) Kelas XI Pada Mata Pelajaran Kimia di SMA Negeri 1 Benail . Penelitian ini diharapkan dapat memberikan kontribusi dalam meningkatkan mutu pembelajaran dan menghasilkan evaluasi yang lebih objektif.

B. Identifikasi Masalah

Berdasarkan latar belakang masalah diatas, identifikasi masalah pada penelitian ini yaitu: Ketidaksesuaian Antara Ekspektasi Guru dan Hasil Ujian Siswa. Meskipun Soal UAS dibuat melalui diskusi sesama guru mata pelajaran kimia dengan harapan meningkatkan kualitas soal, terdapat masalah di mana hasil ujian siswa yang dianggap berprestasi tidak sesuai dengan ekspektasi. Hal ini menunjukkan bahwa mungkin ada perbedaan antara tingkat kesulitan soal yang disusun dengan kemampuan siswa, atau bahwa soal-soal tersebut tidak cukup representatif dalam mengukur pemahaman siswa secara menyeluruh.

⁷ R Amelia, Anggi Ristiyana, dan Sri Rejeki Astuti, —Chemistry learning outcomes assessment: How is the quality of the tests made by the teacher?,|| *Journal of Educational Chemistry (JEC)* 3, no. 1 (2021).

C. Batasan Masalah

Berdasarkan identifikasi masalah yang disebutkan diatas maka peneliti akan membatasi permasalahan agar penelitian yang dilakukan lebih spesifik dan fokus. Permasalahan yang hendak dikaji adalah analisis kualitas butir soal uas (ujian akhir semester) kelas XI pada mata pelajaran kimia di SMAN 1 Benai Tahun Pelajaran 2023/2024 yaitu: validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh soal.

D. Rumusan Masalah

Berdasarkan batasan masalah diatas, maka rumusan masalah dalam penelitian ini adalah Bagaimana kualitas soal kimia ujian akhir semester (UAS) tahun pelajaran 2023/2024 kelas XI di SMA Negeri 1 Benai berdasarkan validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh (distractor)?

E. Tujuan Penelitian

Tujuan penelitian ini adalah untuk menganalisis kualitas soal kimia ujian akhir semester (UAS) tahun pelajaran 2023/2024 kelas XI di SMA Negeri 1 Benai berdasarkan validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh (distractor).

F. Manfaat Penelitian

Hasil penelitian ini diharapkan memiliki manfaat antara lain sebagai berikut:

1. Manfaat Teoritis: Penelitian ini akan menambah wawasan mengenai pentingnya analisis kualitas butir soal ujian dalam konteks pendidikan, terutama di bidang Kimia. Dengan memahami teori-teori dasar tentang validitas, reliabilitas, daya pembeda, dan tingkat kesukaran soal, penelitian ini dapat menjadi referensi dalam pengembangan instrumen evaluasi yang lebih baik di masa depan. Selain itu Penelitian ini dapat menjadi landasan bagi penelitian lebih lanjut mengenai pengembangan metode evaluasi di bidang pendidikan, khususnya dalam analisis soal-soal ujian di berbagai mata pelajaran, sehingga dapat memperbaiki kualitas sistem penilaian di sekolah-sekolah.
2. Manfaat Praktis:
 - a. Bagi Guru: Hasil penelitian ini akan memberikan umpan balik kepada guru mengenai kualitas soal ujian yang telah disusun. Guru dapat memperbaiki soal yang memiliki tingkat kesukaran tidak sesuai atau daya pembeda yang rendah sehingga evaluasi pembelajaran dapat lebih akurat dan adil.
 - b. Bagi Sekolah: Penelitian ini dapat membantu sekolah dalam mengevaluasi efektivitas soal ujian sebagai alat ukur kemampuan siswa. Dengan demikian, sekolah dapat memperbaiki kualitas butir soal untuk ujian-ujian berikutnya agar lebih sesuai dengan standar pendidikan.
 - c. Bagi Siswa: Soal ujian yang berkualitas dapat memberikan gambaran yang lebih tepat mengenai kemampuan dan penguasaan

materi oleh siswa. Hal ini juga dapat membantu siswa dalam mempersiapkan diri untuk ujian dengan lebih baik dan meningkatkan motivasi belajar mereka.

- d. Bagi Peneliti: Penelitian ini memberikan pengalaman langsung dalam menganalisis kualitas butir soal serta mengasah kemampuan peneliti dalam penerapan teori evaluasi pendidikan. Selain itu, penelitian ini juga memperluas wawasan peneliti mengenai penyusunan instrumen evaluasi yang valid dan reliabel.
- e. Bagi Peneliti Selanjutnya: Penelitian ini dapat dijadikan sebagai referensi dan acuan dalam mengembangkan penelitian sejenis, baik dengan objek, metode, maupun variabel yang berbeda. Dengan demikian, peneliti selanjutnya dapat memperluas kajian terkait analisis kualitas butir soal serta menghubungkannya dengan aspek lain, seperti model pembelajaran, karakteristik siswa, maupun kurikulum yang digunakan.

BAB II

TINJAUAN PUSTAKA

A. Kajian Teoritis

1. Pengertian Evaluasi Pembelajaran

Evaluasi pembelajaran adalah suatu proses pengumpulan data untuk menentukan sejauh mana hasil belajar yang dicapai oleh peserta didik setelah mengikuti proses pembelajaran. Evaluasi pembelajaran berfungsi sebagai alat untuk menilai tingkat pencapaian kompetensi siswa dan sebagai bahan refleksi untuk memperbaiki proses belajar mengajar di masa mendatang.⁸ Adapun menurut Sukardi, evaluasi pembelajaran merupakan kegiatan untuk mengumpulkan dan mengolah informasi mengenai proses dan hasil belajar yang diperoleh siswa guna memperoleh gambaran mengenai tingkat pencapaian tujuan pembelajaran. Evaluasi ini meliputi penilaian terhadap ranah kognitif, afektif, dan psikomotor.⁹

Evaluasi pembelajaran, menurut berbagai ahli, pada dasarnya adalah proses pengumpulan dan analisis informasi untuk mengetahui tingkat pencapaian tujuan pembelajaran dan efektivitas suatu program pendidikan.¹⁰ Evaluasi ini dilakukan secara sistematis dan mencakup berbagai aspek belajar siswa, baik pengetahuan, keterampilan, maupun sikap. Dalam pendidikan kimia, evaluasi bertujuan untuk mengukur penguasaan siswa terhadap konsep-konsep kimia,

⁸ Suharsimi Arikunto, *Prosedur penelitian suatu pendekatan praktik* (Jakarta: Bumi Aksara, 2020).

⁹ Sukardi, *Metodologi penelitian pendidikan* (Jakarta: Bumi Aksara, 2016).

¹⁰ Akhiruddin, *Belajar dan pembelajaran* (Kabupaten Gowa: Cahaya Bintang Cemerlang, 2019).

pemahaman fenomena kimia, serta keterampilan dalam menyelesaikan masalah yang berkaitan dengan kimia.

Evaluasi dalam pendidikan dapat dibagi menjadi beberapa jenis berdasarkan tujuan, waktu pelaksanaan, dan sifat pengukurannya.¹¹ Salah satu jenis yang paling umum adalah evaluasi formatif, yang dilakukan sepanjang proses pembelajaran untuk memantau perkembangan siswa dan memberikan umpan balik yang berguna bagi perbaikan pengajaran.¹² Evaluasi formatif bertujuan untuk memperbaiki proses belajar mengajar agar lebih efektif, bukan sekadar untuk menilai hasil akhir.¹³ Contoh evaluasi formatif meliputi kuis harian, tugas rumah, diskusi kelompok, dan pertanyaan reflektif yang membantu guru mengetahui pemahaman siswa secara berkelanjutan, sehingga dapat melakukan penyesuaian pengajaran dengan cepat.

Selain evaluasi formatif, ada juga evaluasi sumatif yang dilaksanakan di akhir suatu periode pembelajaran untuk menilai pencapaian keseluruhan dari tujuan yang telah ditetapkan.¹⁴ Evaluasi sumatif lebih bersifat final dan bertujuan untuk menentukan sejauh mana siswa telah menguasai materi yang diajarkan. Bentuk evaluasi

¹¹ Rohmatus Sholiha dan Maulfi Syaiful Rizal, —Pelaksanaan dan hambatan evaluasi formatif dalam pembelajaran menulis teks laporan hasil observasi di SMK PGRI 3 Malang,‖ *Jurnal Pendidikan Bahasa* 12, no. 1 (1 Juli 2023): 192–209, <https://doi.org/10.31571/bahasa.v12i1.5719>.

¹² F Hidayat, —Analisis kualitas butir soal kimia untuk ujian akhir semester di SMA Negeri 3 Medan,‖ *Jurnal Evaluasi dan Penilaian Pendidikan* 10, no. 2 (2022): 100–110.

¹³ S Pahmi, C Hudaya, dan A Jaya, —Pengaruh model pembelajaran Tefa (Teaching Factory) dalam pembelajaran fisika untuk meningkatkan hasil belajar siswa SMK,‖ *Jurnal Hasil Kajian, Inovasi, dan Aplikasi Pendidikan Fisika* 9, no. 1 (2023): 1–15.

¹⁴ I Magdalena, D Oktavia, dan P Nurjamilah, —Analisis evaluasi sumatif dalam pembelajaran tematik siswa kelas VI SDN Batujaya di era pandemi Covid-19,‖ *Arzusin* 1, no. 1 (2021).

ini sering ditemukan dalam ujian akhir, ujian semester, atau penilaian proyek akhir.¹⁵ Evaluasi sumatif membantu guru dan pihak sekolah dalam memberikan nilai atau peringkat kepada siswa, serta mengevaluasi efektivitas metode pengajaran yang telah diterapkan. Kedua jenis evaluasi ini saling melengkapi, dengan evaluasi formatif lebih berfokus pada proses dan perbaikan, sedangkan evaluasi sumatif lebih menekankan pada hasil akhir dari pembelajaran.¹⁶

2. Konsep Analisis Butir Soal

Di bidang pendidikan, penilaian merupakan bagian yang sangat penting untuk mengukur sejauh mana siswa telah menguasai materi yang diajarkan. Namun, untuk memastikan bahwa penilaian tersebut dapat menggambarkan kemampuan siswa secara akurat, penting untuk memiliki instrumen tes yang berkualitas. Salah satu cara untuk menilai kualitas soal dalam tes adalah melalui analisis butir soal. Analisis butir soal merupakan proses untuk mengevaluasi kualitas setiap item dalam suatu tes, seperti pertanyaan pilihan ganda, guna menilai sejauh mana item tersebut mampu mengukur pengetahuan atau keterampilan yang diharapkan.¹⁷ Proses ini berguna untuk mengidentifikasi pertanyaan yang memerlukan perbaikan, penghapusan, atau penyesuaian agar lebih selaras dengan tujuan

¹⁵ I Magdalena, D Oktavia, dan P Nurjamilah, —Analisis evaluasi sumatif dalam pembelajaran tematik siswa kelas VI SDN Batujaya di era pandemi Covid-19,|| *Arzusin* 1, no. 1 (2021).

¹⁶ L Purniasari et al., —Analisis butir soal ujian sekolah mata pelajaran kimia SMA N 1 Kutowinangun tahun pelajaran 2019/2020 menggunakan model Iteman dan Rasch,|| *Jurnal Pendidikan Kimia* 10, no. 2 (2021): 205–14.

¹⁷ Rizka Putri Altı, Zulyuri, dan Violita, —Analisis kualitas butir soal ujian tengah semester mata pelajaran biologi kelas X di MAN 1 Solok Selatan,|| *Jurnal Metaedukasi* 4, no. 2 (2022).

pembelajaran. Dengan demikian, analisis ini berperan penting dalam memastikan bahwa penilaian dilakukan secara objektif dan akurat.¹⁸

Kualitas butir soal merujuk pada seberapa baik suatu butir soal dapat memenuhi fungsi sebagai alat ukur dalam evaluasi pembelajaran.¹⁹ Kualitas butir soal ditentukan oleh beberapa indikator dalam menganalisis kualitas butir soal sebagai berikut:

a. Daya Pembeda

Daya pembeda adalah kemampuan suatu butir soal untuk membedakan antara siswa yang memiliki kemampuan tinggi dan rendah. Soal yang memiliki daya pembeda tinggi dapat mengidentifikasi siswa yang memahami materi dengan baik. Menurut penelitian, daya pembeda yang optimal adalah di atas 0,20, dan soal yang tidak memenuhi kriteria ini sebaiknya direvisi atau dihilangkan dari tes.²⁰

b. Tingkat kesukaran

Tingkat kesukaran soal merujuk pada persentase siswa yang dapat menjawab suatu soal dengan benar. Soal dengan tingkat kesukaran terlalu tinggi atau terlalu rendah dianggap tidak efektif dalam mengukur kemampuan siswa. Idealnya, tingkat kesukaran berada di antara 0,30 hingga 0,70. Penelitian oleh Anggraeni

¹⁸ A Sabara et al., —Validitas perangkat pembelajaran berbasis STEAM materi struktur dan fungsi jaringan tumbuhan untuk meningkatkan penguasaan konsep,|| *Jambura Edu Biosfer Journal* 4, no. 2 (2022): 76–85.

¹⁹ H D Ikawati, A Jayadi, dan Hermyans, —Analisis kualitas tes dan butir soal sejarah di SMAN 1 Praya Timur,|| *Jurnal Pendidikan dan Konseling* 4 (2022): 1349–58.

²⁰ S Wulandari, —Pengaruh daya pembeda terhadap kualitas soal pilihan ganda di sekolah menengah,|| *Jurnal Penilaian Pendidikan* 10, no. 1 (2022): 78–86.

(2019) menemukan bahwa sebagian besar soal yang dianalisis memiliki tingkat kesukaran yang tidak sesuai dengan standar ideal, yang mengarah pada revisi soal untuk mencapai tingkat kesulitan yang seimbang.²¹

c. Reliabilitas Butir Soal

Reliabilitas menunjukkan konsistensi hasil yang diberikan oleh suatu tes. Soal yang reliabel akan menghasilkan skor yang relatif sama ketika diujikan dalam kondisi yang sama. Pengujian reliabilitas dapat dilakukan dengan berbagai metode, seperti koefisien *Cronbach's Alpha* atau *Kuder-Richardson*.²²

d. Validitas Butir Soal

Validitas adalah kemampuan suatu butir soal untuk mengukur apa yang seharusnya diukur. Validitas dapat dilihat dari kesesuaian butir soal dengan tujuan pembelajaran yang telah ditetapkan. Dalam analisis kualitas butir soal, validitas dapat diuji secara empiris melalui korelasi antara skor butir soal dengan skor total tes.²³

e. Efektivitas pengecoh

Efektivitas pengecoh mengacu pada kemampuan opsi jawaban yang salah dalam soal pilihan ganda untuk menarik perhatian siswa yang tidak memahami materi. Pengecoh yang

²¹ Anggraeni, *Op. Cit.*

²² D Rahmawati dan N Yuliani, —Peningkatan reliabilitas butir soal melalui analisis empiris,| *Jurnal Evaluasi Pendidikan* 12, no. 2 (2020): 110–20.

²³ Wibowo, *Op. Cit.*

efektif adalah yang dipilih oleh sebagian siswa, menunjukkan bahwa mereka belum memahami materi tersebut dengan benar.²⁴

3. Pembelajaran Kimia

Pembelajaran adalah suatu usaha yang dilakukan secara sadar dan terencana, di mana siswa harus memahami tujuan yang telah ditentukan sebelum proses dimulai. Proses ini dilaksanakan dengan pengendalian yang baik terhadap aspek-aspek seperti isi, waktu, proses, dan hasil. Pembelajaran merupakan kombinasi dari berbagai komponen yang saling mempengaruhi dalam proses untuk mencapai tujuan pembelajaran. Hal ini mencakup proses perubahan perilaku, bukan sekadar hasil akhir yang dicapai. Dalam pembelajaran, terdapat dua subjek yang terlibat, yaitu guru dan siswa, yang kemudian menghasilkan perubahan pada diri siswa sebagai hasil dari proses tersebut.²⁵

Proses pembelajaran melibatkan dua pihak, yaitu guru dan siswa, yang berkontribusi pada perubahan pada siswa sebagai hasil dari kegiatan belajar. Berdasarkan UU No.23 Tahun 2020 tentang sistem pendidikan nasional, pembelajaran didefinisikan sebagai proses interaksi antara siswa dengan guru serta berbagai sumber belajar

²⁴ A Santoso dan M Dewi, —Efektivitas pengecoh pada soal ujian pilihan ganda,|| *Jurnal Ilmu Pendidikan* 21, no. 3 (2023): 30–40.

²⁵ B I Sappaile dan T Pristiwaluyo, —Analisis butir soal ujian sekolah berstandar nasional dengan pendekatan klasik dan teori respon butir mata pelajaran matematika,|| in *Proceedings of National Seminar Research and Community Service Institue Universitas Negeri Makassar*, vol. 15, 2019, 246–53.

dalam lingkungan tertentu.²⁶ Dalam pembelajaran kimia, teori belajar yang relevan adalah teori belajar penemuan yang dikemukakan oleh Bruner dan teori belajar bermakna oleh Ausubel. Oleh karena itu, dalam mengimplementasikan pembelajaran kimia, perlu dirancang sistem pembelajaran yang sesuai dengan karakteristik ilmu kimia, di mana pengetahuan diperoleh melalui proses dan metode ilmiah.²⁷ Dengan demikian, pembelajaran kimia dapat disimpulkan sebagai kegiatan belajar yang melibatkan guru dan siswa dalam konteks mata pelajaran kimia, baik dalam bentuk teori maupun praktikum.

Tujuan pembelajaran kimia adalah untuk mendapatkan pemahaman yang mendalam mengenai berbagai fakta, mampu mengenali dan menyelesaikan masalah, memiliki keterampilan dalam penggunaan laboratorium, serta mengembangkan kemandirian belajar yang dapat diterapkan dalam kehidupan sehari-hari.²⁸ Bidang kimia mencakup area yang luas, dengan fakta, konsep, dan teorinya yang terus berkembang. Sebagian materi dapat diajarkan melalui metode penemuan, inkuiri, dan pemecahan masalah, sementara sebagian besar diajarkan melalui pembelajaran konvensional seperti ceramah dan tugas membaca mandiri. Proses pembelajaran perlu dilakukan secara inovatif agar siswa tidak merasa bosan. Inovasi dalam pembelajaran

²⁶ Republik Indonesia, *Undang-Undang Republik Indonesia Nomor 23 Tahun 2020 tentang Sistem Pendidikan Nasional* (Jakarta: Pemerintah Pusat, 2020).

²⁷ Sappaile dan Pristiwaluyo, *Op. Cit.*

²⁸ Wibowo, *Op. Cit.*

kimia dapat meningkatkan motivasi dan prestasi siswa.²⁹ Materi pembelajaran untuk kelas XI pada semester 2 meliputi Larutan Elektrolit dan Asam Basa, Hidrolisis Garam, Larutan Penyangga, Kelarutan dan Hasil Kali Kelarutan, serta Koloid yang keseluruhannya dapat dilihat pada lampiran 4.

Materi Larutan Elektrolit dan Asam Basa membahas tentang sifat larutan yang mengandung zat yang dapat mengion dalam air.³⁰ Zat yang dapat mengion disebut elektrolit, yang dibedakan menjadi elektrolit kuat (seperti NaCl, HCl) dan elektrolit lemah (seperti asam asetat). Elektrolit kuat dapat terionisasi sepenuhnya dalam air, sedangkan elektrolit lemah hanya terionisasi sebagian. Selain itu, pembahasan juga meliputi konsep asam basa berdasarkan teori Arrhenius, yang menyatakan bahwa asam melepaskan ion H^+ dan basa melepaskan ion OH^- dalam air. Pemahaman tentang pH dan pengukuran konsentrasi ion H^+ atau OH^- juga menjadi fokus materi ini untuk menjelaskan kekuatan asam dan basa dalam larutan.³¹

Hidrolisis Garam adalah proses reaksi antara garam dan air, yang menghasilkan ion-ion yang dapat mempengaruhi pH larutan.³² Garam yang terbentuk dari asam kuat dan basa kuat tidak mengalami hidrolisis, sehingga larutannya bersifat netral. Namun, garam yang

²⁹ F Amalia dan T Nugroho, —Evaluasi daya pembeda soal dalam pembelajaran kimia, *Jurnal Pendidikan dan Pembelajaran Kimia* 10, no. 2 (2022): 100–110.

³⁰ Kemendikbud RI, *Kimia untuk SMA/MA kelas XI* (Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2014).

³¹ *Ibid.*

³² *Ibid.*

terbentuk dari asam lemah dan basa kuat, atau basa lemah dan asam kuat, dapat mengalami hidrolisis, menghasilkan larutan yang bersifat asam atau basa. Materi ini membahas bagaimana reaksi hidrolisis mempengaruhi pH larutan dan bagaimana perhitungan konstanta hidrolisis digunakan untuk menentukan sifat pH larutan garam tertentu. Pemahaman ini penting dalam konteks kimia analitik dan aplikasi dalam industri, seperti pengolahan air dan pertanian.

Larutan Penyangga adalah larutan yang mampu mempertahankan pH-nya meskipun ditambahkan asam atau basa.³³ Larutan penyangga terdiri dari pasangan asam lemah dan basa konjugatnya atau basa lemah dan asam konjugatnya. Materi ini mengajarkan bagaimana larutan penyangga bekerja untuk mengatur pH, yang sangat penting dalam proses biologis, seperti dalam darah manusia. Selanjutnya, Kelarutan dan Hasil Kali Kelarutan membahas konsep kelarutan zat dalam air dan bagaimana faktor-faktor seperti suhu mempengaruhi kelarutan.³⁴ Hasil kali kelarutan (K_{sp}) digunakan untuk menghitung kelarutan garam yang tidak larut dalam air. Materi ini penting untuk memahami bagaimana zat-zat terlarut dan tak terlarut berinteraksi dalam larutan. Terakhir, Koloid adalah sistem campuran di mana partikel terdispersi (seperti protein atau partikel nano) tersebar dalam fase lain (seperti air), yang tidak dapat

³³ Kemendikbud RI, *Kimia untuk SMA/MA kelas XI* (Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2014).

³⁴ *Ibid.*

dipisahkan dengan penyaringan biasa.³⁵ Koloid memiliki sifat unik, seperti efek Tyndall, yang memungkinkan partikel-partikelnya tersebar di seluruh larutan, dan digunakan dalam berbagai aplikasi, seperti dalam pengemulsi dan pengolahan makanan.

B. Penelitian Relevan

Beberapa penelitian yang relevan dalam penelitian ini adalah:

1. Hasil penelitian dari Anis Fauziana dan Andhita Dessy Wulansari, yang berjudul: "Analisis Kualitas Butir Soal Ulangan Harian di Sekolah Dasar dengan Model Rasch", semua soal yang diujikan pada peserta tes adalah valid dan reliabilitas. Hasil Tingkat kesukaran dari 10 soal yang diuji adalah 4 soal tergolong mudah, 4 soal tergolong sedang dan 2 soal tergolong sangat sukar. Berdasarkan penelitian yang telah dilakukan dapat juga diketahui bahwa nilai estimasi kemampuan orang dari 31 siswa bervariasi pada rentang -0.69 sampai dengan 3,05. Hal ini menunjukkan kemampuan siswa tergolong sedang sampai sangat tinggi.³⁶
2. Hasil penelitian dari Leonardus Kaka, dkk., yang berjudul: "Efektivitas Analisis Butir Soal Pilihan Ganda Menggunakan Aplikasi Anates di SMPN 2 Kanatang", Hasil penelitian menunjukkan butir soal yang dipakai tergolong sukar 55%, butir soal yang tergolong sedang 45% dan butir soal yang tergolong sukar sekali 5%. Analisis

³⁵ *Ibid.*

³⁶ A Fauziana dan Andhita Dessy Wulansari, —Analisis kualitas butir soal ulangan harian di sekolah dasar dengan model Rasch,|| *Jurnal Kependidikan Dasar Islam Berbasis Sains* 6, no. 1 (2021).

daya beda soal menunjukkan 30% butir soal tersebut sudah memiliki daya beda yang baik, soal yang memiliki daya pembeda sedang 28% sedangkan soal yang daya pembedanya tidak baik 42% perlu direvisi. Analisis efektivitas *option* pengecoh dari butir soal diketahui 88% dari total *option* pengecoh pada butir soal tersebut termasuk baik dan sisanya tidak baik dan perlu direvisi. Berdasarkan hasil analisis tingkat validitas di kategorikan soal yang baik 26 (88%) sedangkan soal yang tidak valid 4 (12%) dibuang. Koefisien reliabilitas sebesar 0.62 sehingga disebut soal belum memiliki reliabilitas yang tinggi atau reliabel.³⁷

3. Penelitian oleh Mustaqim & Hidayu Sulisti, yang berjudul: "Analisis Butir Soal Pas Matematika Peminatan: Daya Pembeda, Tingkat Kesukaran, Dan Kualitas Pengecoh", menunjukkan bahwa, daya pembeda soal PAS genap mata pelajaran matematika peminatan kelas XII IPA 1 SMA Negeri 10 Pontianak paling banyak berada pada kategori cukup. Hal ini artinya kualitas soal tersebut mampu membedakan dengan cukup baik peserta didik yang sudah memahami dan yang belum. Untuk tingkat kesukaran, paling banyak berada pada kategori mudah. Hal ini berarti soal kurang dapat merangsang siswa untuk mempertinggi usaha dalam memecahkan suatu permasalahan yang diberikan. Sedangkan untuk kualitas pengecoh, berada pada kategori sangat baik sebesar 22%, kategori baik sebesar 11,5%,

³⁷ Leonardus Kaka, Oktoviana Bano, dan Yohana Njoeroemana, —Efektivitas analisis butir soal pilihan ganda menggunakan aplikasi anates di SMPN 2 Kanatang,|| *Jurnal Inovasi Penelitian* 4, no. 9 (2024).

kurang baik 13%, buruk 29%, dan kualitas pengecoh pada kategori sangat buruk sebesar 24,5%.³⁸

4. Penelitian oleh Nurul Hidayah Siregar, dkk., yang berjudul: "Analisis Butir Soal Ujian Tengah Semester Pada Mata Pelajaran Pendidikan Agama Islam", Hasil penelitian analisis soal ujian semester menunjukkan bahwa validitas yang ditemukan pada analisis soal sebanyak 15 pada taraf 75% sedangkan yang tidak valid sebanyak 5 orang pada taraf 5%. Reliabilitas nya ialah sebesar 0,775. Maka dapat disimpulkan bahwa data dari butir soal memiliki kriteria yang tinggi. Indek kesukaran yan di dapat yaitu pada kriteria sedang sebanyak 12 soal pada persentase 60 %. Daya beda soal analisis nya pada kategori baik dengan jumlah soal sebanyak 8 pada persentase 40 %. Dengan demikian setiap soal yang dibuat oleh pendidik diharapkan memiliki tingkat validitas, reliabilitas, indeks kesukaran soal, serta daya beda yang baik sesuai dengan kategori atau kriteria soal yang dibuat.³⁹
5. Penelitian oleh Hendrik Simamora, dkk., yang berjudul: "Analisis Kualitas Butir Soal Buatan Guru Kimia Pada Tes Ujian Tengah Semester Ganjil Kelas XII MIPA", Hasil dari penelitian ini disimpulkan dengan hasil dari analisis yang memiliki berbagai kategori seperti 2 butir (10%) untuk sangat baik dan 2 butir (10%) juga dengan kategori sangat tidak baik dan disarankan untuk dibuang,

³⁸ Mustaqim dan Hidayu Sulisti, —Analisis butir soal PAS matematika peminatan: Daya pembeda, tingkat kesukaran, dan kualitas pengecoh,|| *Al- 'Adad: Jurnal Tadris Matematika* 3, no. 1 (2024).

³⁹ Nurul Hidayah Siregar, Remiswal, dan Khadijah, —Analisis butir soal ujian tengah semester pada mata pelajaran pendidikan agama Islam,|| *Urwatul Wutsqo: Jurnal Studi Kependidikan dan Keislaman* 13, no. 2 (2024).

8 butir soal baik, 4 butir soal cukup, 3 butir soal tidak baik dan disimpulkan pada penelitian ini kualitas butir soal yang diujikan adalah sedang.⁴⁰

Berikut ini tabel penelitian relevan:

Tabel 2.1 Tabel Penelitian Relevan

No.	Nama Peneliti	Mata Pelajaran	Kelas	Jumlah Butir Soal	Jenis Tes	Aplikasi Analisis Data
1.	Anis dan Andhita	Matematika	VI	20 Butir	Formatif	Rasch
2.	Leonardus K., dkk.	IPA	VII	30 Butir	Sumatif	Anates
3.	Mustaqim dan Hidayu Sulisti	Matematika	XII	40 Butir	Sumatif	Anates
4.	Nurul H. S., dkk.	PAI	XII	20 Butir	Sumatif	SPSS
5.	Hendrik Simamora, dkk.	Kimia	XII	20 Butir	Sumatif	SPSS dan Anates
6.	Maya Febriani C.	Kimia	XI	25 Butir	Sumatif	Iteman

C. Kerangka Konseptual

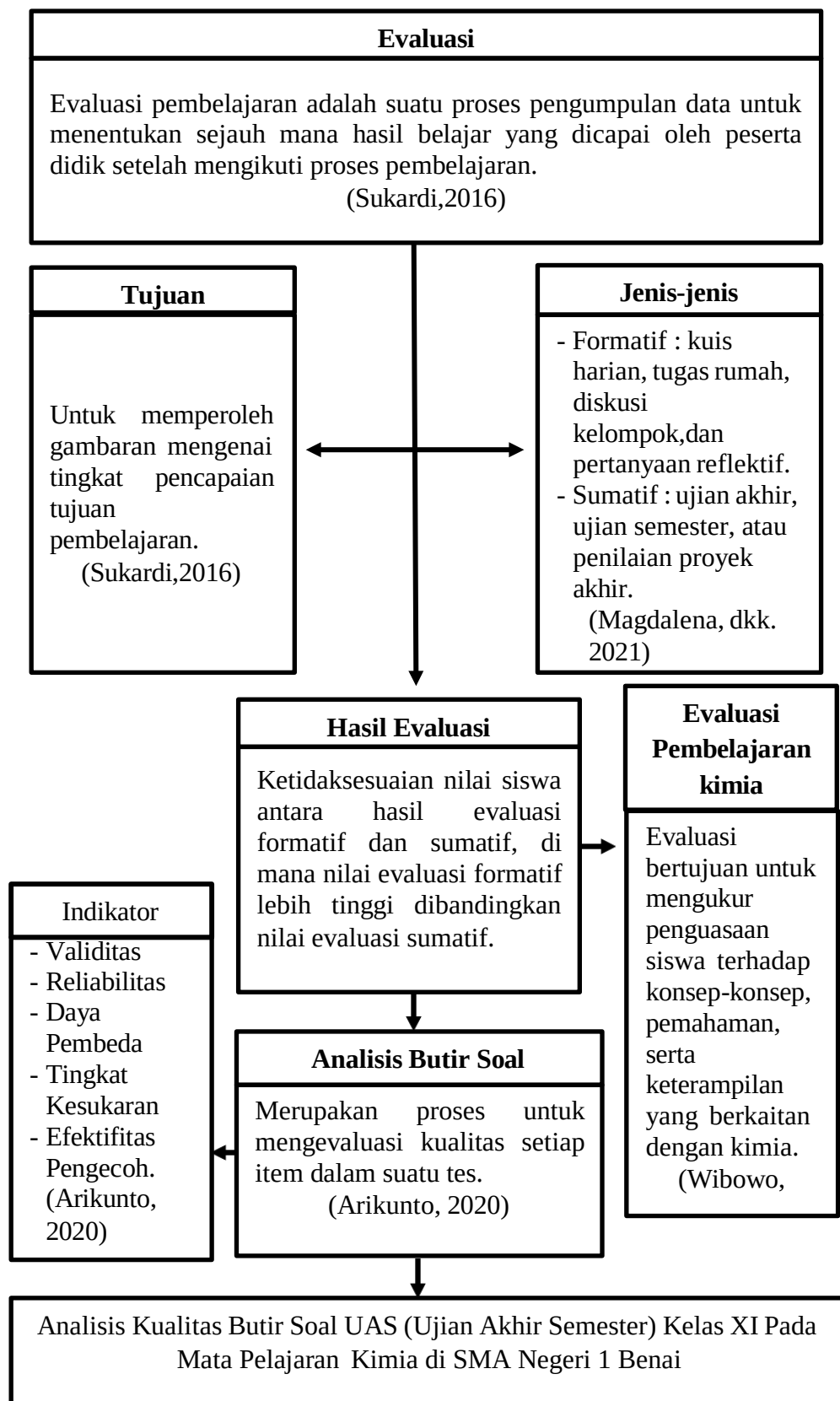
Tujuan pengajaran yang hendak dicapai di sekolah mempunyai kaitan dengan materi dan metode belajar mengajar yang dipakai guru dalam memberikan materi, dan siswa dalam menerima materi tersebut. Sejah mana keberhasilan guru memberikan materi dan sejah mana siswa menyerap materi yang disajikan itu dapat diperoleh informasinya

⁴⁰ Hendrik Simamora, Hartono, dan Effendi, —Analisis kualitas butir soal buatan guru kimia pada tes ujian tengah semester ganjil kelas XII MIPA,|| *Hydrogen: Jurnal Kependidikan Kimia* 9, no. 1 (2021).

melalui evaluasi. Analisis butir soal Ujian Akhir Semester (UAS) digunakan oleh guru untuk mengetahui bagaimana peserta didiknya dalam menerima pelajaran selama ini. Dalam melakukan analisis butir soal perlu diperhatikan beberapa hal, yaitu sebagai berikut:

1. Validitas
2. Reliabilitas
3. Tingkat kesukaran soal
4. Daya pembeda
5. Efektivitas pengecoh

Oleh karena itu diperlukan analisis butir soal yang digunakan pada UAS (Ujian Akhir Semester) Kelas XI Pada Mata Pelajaran Kimia di SMAN 1 Benai. Dimana dengan hasil analisis butir soal tersebut akan diketahui soal mana yang berkualitas dan tidak berkualitas. Gambaran lebih jelas mengenai kerangka konseptual dapat dilihat pada bagan berikut ini:



Gambar 2.1. Bagan Kerangka Konseptual

BAB III

METODOLOGI PENELITIAN

A. Jenis Penelitian

Penelitian ini merupakan penelitian kuantitatif deskriptif, yang bertujuan untuk mendeskripsikan dan menganalisis kualitas butir soal Ujian Akhir Semester (UAS) mata pelajaran Kimia kelas XI di SMA Negeri 1 Benai. Menurut Sugiyono, penelitian kuantitatif deskriptif adalah suatu penelitian yang digunakan untuk menggambarkan atau menjelaskan suatu fenomena dengan menggunakan data berupa angka-angka yang diolah secara statistik. Tujuan utama dari penelitian ini adalah untuk memberikan gambaran secara sistematis, faktual, dan akurat mengenai fakta-fakta serta sifat-sifat populasi atau objek tertentu.⁴¹ Penelitian kuantitatif deskriptif dipilih karena sesuai untuk mengukur dan menganalisis data yang bersifat numerik serta memberikan deskripsi tentang kualitas butir soal UAS secara objektif. Dengan pendekatan ini, hasil penelitian diharapkan dapat memberikan informasi yang jelas dan mendetail mengenai kekuatan dan kelemahan dari butir soal yang digunakan dalam UAS Kimia, serta rekomendasi untuk perbaikan kualitas soal.

B. Waktu dan Lokasi Penelitian

1. Waktu penelitian: Penelitian ini dilaksanakan pada bulan Oktober 2024 sampai bulan April 2025.

⁴¹ Sugiyono, *Metode penelitian kuantitatif, kualitatif, dan R&D* (Bandung: Alfabeta, 2017).

2. Lokasi penelitian: Penelitian dilaksanakan di SMA Negeri 1 Benai, yang berlokasi di Kabupaten Kuantan Singingi, Riau. Sekolah ini dipilih karena terdapat penggunaan soal UAS Kimia yang menjadi subjek analisis dalam penelitian ini.

C. Populasi dan Sampel

A. Populasi

Populasi dalam penelitian ini adalah siswa kelas XI yang memilih mata pelajaran kimia sebagai mata pelajaran pilihan, yaitu, kelas XI.1, XI.3, XI.4, dan XI.6 yang berjumlah 144 orang.

B. Sampel

Sampel sebanyak 106 siswa menggunakan metode *simple random sampling* dengan *margin of error* 5%. Dengan teknik ini, setiap siswa di dalam populasi memiliki peluang yang sama untuk terpilih menjadi bagian dari sampel, tanpa ada diskriminasi berdasarkan jurusan, nilai, atau karakteristik lainnya.⁴² Teknik *random sampling* ini digunakan untuk memastikan bahwa sampel yang diambil dapat mewakili populasi secara keseluruhan.

D. Teknik Pengumpulan Data

Dalam penelitian ini, teknik pengumpulan data yang digunakan adalah data sekunder (data yang sudah ada) sebagai sumber data utama, yang diperoleh dari hasil ujian akhir semester kelas XI Tahun Ajaran 2023/2024 pada mata pelajaran Kimia.

⁴² Agung Santoso, —Rumus Slovin: Panacea masalah ukuran sampel?,*l Suksma: Jurnal Psikologi Universitas Sanata* 4, no. 2 (2023).

E. Teknik Analisis Data

Data sekunder ini dianalisis menggunakan teknik uji statistik yang sesuai untuk mendapatkan informasi yang diinginkan dari data yang sudah matang. Analisis ini bertujuan untuk mengukur kualitas butir soal Ujian Akhir Semester (UAS) dengan menggunakan indikator-indikator tertentu, seperti validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Teknik analisis yang digunakan mencakup beberapa tahap sebagai berikut:

a. Analisis Validitas Butir Soal

Validitas butir soal mengukur sejauh mana butir soal dapat mengukur apa yang seharusnya diukur. Validitas isi dinilai dengan melihat kesesuaian soal dengan indikator pembelajaran dan materi yang diajarkan. Untuk menganalisis validitas empiris, digunakan korelasi antara skor butir soal dengan skor total tes. Validitas dapat diuji dengan menggunakan teknik korelasi point biserial.⁴³

Korelasi point biserial adalah jenis korelasi khusus yang mengukur hubungan antara variabel dikotomi (seperti benar atau salah) dan variabel kontinu (seperti skor total). Berbeda dengan korelasi biserial, point biserial tidak mengasumsikan bahwa variabel dikotomi berasal dari distribusi kontinu yang mendasari. Korelasi ini

⁴³ Arikunto, *Loc. Cit*, 2020.

cocok digunakan saat variabel dikotomi benar-benar diskret.⁴⁴ Rumus

Korelasi Point Biserial:⁴⁵

$$r_{pb} = \frac{(M_1 - M_0)}{S} \times \sqrt{\frac{p_1 p_0}{N}}$$

Di mana:

M_1 = rata-rata skor total siswa yang menjawab benar

M_0 = rata-rata skor total siswa yang menjawab salah

S = standar deviasi skor total

p_1 = proporsi siswa yang menjawab benar

p_0 = proporsi siswa yang menjawab salah ($1 - p_1$)

N = jumlah total siswa

Kriteria validitas butir soal:⁴⁶

- $\geq 0,40$ = Validitas sangat tinggi
- $0,30 - 0,39$ = Validitas tinggi
- $0,20 - 0,29$ = Validitas sedang
- $0,00 - 0,19$ = Validitas rendah
- $< 0,00$ = Tidak valid

Korelasi point biserial mengukur kekuatan dan arah hubungan antara jawaban siswa pada suatu butir soal dan skor total mereka. Semakin tinggi nilai korelasi, semakin baik soal tersebut dalam membedakan siswa dengan tingkat kemampuan yang berbeda.

⁴⁴ Arikunto, *Loc. Cit.*, 2020.

⁴⁵ *Ibid.*

⁴⁶ *Ibid.*

b. Analisis Reliabilitas Butir Soal

Reliabilitas mengukur konsistensi atau kestabilan hasil tes. Dalam penelitian ini, reliabilitas dihitung menggunakan koefisien *Alpha Cronbach* untuk soal pilihan ganda atau rangkaian tes ulang (*test-retest*) untuk melihat konsistensi hasil jika tes diberikan kembali pada waktu yang berbeda.

Koefisien *Alpha Cronbach* merupakan salah satu metode yang paling umum digunakan untuk mengukur reliabilitas, terutama untuk soal tes pilihan ganda. Nilai *Alpha Cronbach* menunjukkan seberapa konsisten item-item dalam tes dalam mengukur konstruk yang sama. Rumus untuk menghitung *Alpha Cronbach* adalah sebagai berikut:⁴⁷

$$\alpha = \frac{N}{N-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right)$$

Di mana:

N adalah jumlah butir soal

σ_i^2 adalah varians skor setiap butir soal

σ_t^2 adalah varians total dari skor tes

Nilai *Alpha Cronbach* berkisar antara 0 hingga 1. Semakin tinggi nilainya (mendekati 1), semakin tinggi konsistensi internal item-item dalam tes tersebut.⁴⁸ Sebagai panduan, nilai reliabilitas dapat dikategorikan sebagai berikut:⁴⁹

⁴⁷ Arikunto, *Loc. Cit*, 2020.

⁴⁸ *Ibid.*

⁴⁹ *Ibid.*

1) Sangat Baik ($\geq 0,90$)

- Tes memiliki tingkat reliabilitas yang sangat tinggi.
- Soal sangat baik dan layak digunakan tanpa revisi.

2) Baik ($0,80-0,89$)

- Tes memiliki reliabilitas tinggi.
- Soal baik dan layak digunakan, revisi kecil mungkin diperlukan.

3) Cukup ($0,70-0,79$)

- Tes memiliki reliabilitas cukup baik.
- Soal layak digunakan, namun perlu perbaikan untuk meningkatkan kualitas.

4) Kurang Baik ($0,60-0,69$)

- Tes memiliki reliabilitas yang cukup rendah.
- Soal dapat digunakan dengan catatan harus dilakukan revisi besar.

5) Tidak Baik ($< 0,60$)

- Tes memiliki reliabilitas yang sangat rendah.
- Soal tidak layak digunakan dan perlu revisi menyeluruh atau diganti.

c. Analisis Daya Pembeda

Daya pembeda mengukur kemampuan suatu soal untuk membedakan antara siswa yang memiliki kemampuan tinggi dan rendah. Rumus yang digunakan untuk menghitung daya pembeda adalah:⁵⁰

⁵⁰Arikunto, *Loc. Cit*, 2020.

$$D = \frac{(\text{Jumlah Siswa Kelompok Atas Yang Benar}) - (\text{Jumlah Siswa Kelompok Kecil Yang Benar})}{\text{Jumlah Siswa Dalam Satu Kelompok}}$$

Kriteria daya pembeda dapat dikategorikan sebagai berikut:

- 1) Baik: jika $D > 0,40$, soal memiliki daya pembeda yang baik dan efektif membedakan antara siswa berkemampuan tinggi dan rendah.
- 2) Cukup: jika $0,20 < D \leq 0,40$, soal memiliki daya pembeda yang cukup dan masih dapat digunakan dengan perbaikan.
- 3) Kurang: jika $D \leq 0,20$, soal kurang mampu membedakan kemampuan siswa dan sebaiknya diperbaiki.
- 4) Dibuang/diganti: jika $D < 0,00$ (negatif), soal dianggap tidak efektif, karena menunjukkan bahwa siswa berkemampuan rendah lebih banyak menjawab benar dibanding siswa berkemampuan tinggi, sehingga perlu dibuang atau diganti.

d. Analisis Tingkat Kesukaran

Tingkat kesukaran butir soal menunjukkan seberapa mudah atau sulit suatu soal bagi siswa. Tingkat kesukaran dihitung dengan menggunakan rumus:⁵¹

$$P = \frac{\text{Jumlah Siswa Yang Menjawab Benar}}{\text{Jumlah Seluruh Siswa}}$$

⁵¹ Arikunto, *Loc. Cit.*, 2020.

Kriteria tingkat kesukaran dibagi menjadi tiga kategori:

- 1) Mudah: $P > 0,70$ (Soal tergolong mudah jika lebih dari 70% siswa menjawab benar)
- 2) Sedang: $0,30 \leq P \leq 0,70$ (Soal tergolong sedang jika 30-70% siswa menjawab benar)
- 3) Sukar: $P < 0,30$ (Soal tergolong sukar jika kurang dari 30% siswa menjawab benar)

e. Analisis Efektivitas Pengecoh

Pengecoh dalam soal pilihan ganda digunakan untuk mengecoh atau "memancing" siswa yang tidak memahami konsep dengan baik. Jika pengecoh tersebut dipilih oleh siswa dalam jumlah yang cukup, maka dapat dianggap sebagai pengecoh yang efektif. Sebaliknya, jika pengecoh tidak pernah dipilih oleh siswa, maka pengecoh tersebut dianggap tidak berfungsi dan tidak efektif. Pengecoh dianggap efektif apabila memenuhi beberapa kriteria berikut:⁵²

- 1) Daya tarik terhadap siswa: Pengecoh yang baik dipilih oleh setidaknya 5% dari total siswa yang mengikuti tes. Ini menunjukkan bahwa pengecoh cukup membingungkan bagi sebagian siswa, sehingga berfungsi sesuai dengan tujuannya.
- 2) Distribusi pilihan jawaban: Pengecoh yang baik akan menyebabkan distribusi jawaban yang merata, di mana pengecoh

⁵² Arikunto, *Loc. Cit.*, 2020.

yang kurang efektif hanya dipilih oleh sedikit siswa atau tidak ada sama sekali.

- 3) Menyesatkan siswa yang tidak menguasai materi: Pengecoh yang efektif seharusnya lebih banyak dipilih oleh siswa dengan kemampuan rendah dibandingkan dengan siswa yang berprestasi tinggi.

Untuk menganalisis efektivitas pengecoh, dilakukan langkah-langkah sebagai berikut:⁵³

- 1) Mengumpulkan data jawaban siswa pada setiap butir soal, termasuk pilihan pengecoh.
- 2) Menghitung persentase siswa yang memilih setiap pengecoh.
- 3) Menentukan efektivitas pengecoh berdasarkan kriteria.

Dengan melakukan analisis efektivitas pengecoh, peneliti dapat:

- 1) Mengidentifikasi pengecoh mana yang perlu diperbaiki atau diganti untuk meningkatkan kualitas soal.
- 2) Meningkatkan kualitas evaluasi melalui perbaikan butir soal sehingga dapat mengukur kemampuan siswa secara lebih akurat.
- 3) Menyusun soal pilihan ganda yang lebih valid dan *reliabel*.

Selain itu, teknik analisis data dalam penelitian ini juga menggunakan aplikasi ITEMAN untuk menganalisis kualitas butir soal berdasarkan parameter teori klasik. Data yang dianalisis

⁵³ Suharsimi Arikunto, *Dasar-dasar evaluasi pendidikan*, Revisi (Jakarta: Bumi Aksara, 2019).

mencakup hasil ujian akhir semester (UAS) mata pelajaran kimia kelas XI di SMA Negeri 1 Benai. Dengan ITEMAN, analisis dilakukan untuk menghitung daya pembeda, tingkat kesukaran, dan validitas internal butir soal. Selain itu, aplikasi ini juga digunakan untuk mengidentifikasi pola jawaban siswa melalui analisis opsi jawaban (distraktor), guna memastikan bahwa setiap opsi berfungsi sebagaimana mestinya. Data yang diinputkan adalah hasil pilihan ganda yang telah dikodekan dalam format tertentu, kemudian diproses menggunakan perangkat lunak untuk menghasilkan laporan statistik.

Hasil analisis dari ITEMAN akan digunakan untuk mengevaluasi kualitas keseluruhan soal yang diujikan. Parameter daya pembeda akan menunjukkan sejauh mana butir soal mampu membedakan siswa dengan kemampuan tinggi dan rendah.⁵⁴ Tingkat kesukaran memberikan gambaran proporsi siswa yang menjawab benar setiap butir soal, sehingga butir dapat dikategorikan mudah, sedang, atau sulit.⁵⁵ Sementara itu, efektivitas distraktor akan dianalisis untuk memastikan bahwa semua pilihan jawaban, kecuali kunci jawaban, mampu mengalihkan perhatian siswa yang belum memahami materi dengan baik.⁵⁶ Data ini akan digunakan sebagai dasar untuk merevisi atau menyusun soal yang lebih baik pada periode ujian berikutnya.

⁵⁴ S Azwar, *Reliabilitas dan validitas instrumen penelitian (Edisi Revisi)* (Yogyakarta: Pustaka Pelajar, 2023).

⁵⁵ Suharsimi Arikunto, *Dasar-dasar evaluasi pendidikan* (Jakarta: Bumi Aksara, 2018).

⁵⁶ *Ibid.*

Output ITEMAN merupakan alat analisis butir soal berbasis komputer yang menyajikan sejumlah indikator statistik penting untuk menilai kualitas instrumen tes secara empiris. Program ini menghasilkan parameter seperti *prop. correct* untuk mengukur tingkat kesukaran, *point biserial* untuk menilai daya pembeda, serta *prop. endorsing* untuk mengevaluasi efektivitas pengecoh. Salah satu keunggulan ITEMAN adalah kemampuannya menampilkan indikator validitas dan daya pembeda pada kolom output yang sama, yaitu *point biserial*. Dalam konteks ini, nilai *point biserial* tidak hanya menunjukkan kemampuan butir soal dalam membedakan antara siswa berkemampuan tinggi dan rendah, tetapi juga merefleksikan validitas empiris dari soal tersebut. Artinya, semakin tinggi nilai *point biserial*, semakin kuat hubungan antara jawaban benar terhadap butir soal dengan total skor, sehingga butir soal tersebut dianggap valid dan memiliki daya pembeda yang baik.⁵⁷

Selain itu, *prop. endorsing* memberikan gambaran tentang sebaran pilihan jawaban yang dipilih peserta, termasuk kunci dan pengecoh. Distraktor atau pengecoh yang efektif ditunjukkan dengan nilai proporsi minimal $\geq 5\%$, menandakan bahwa opsi tersebut mampu "menjebak" siswa yang belum menguasai materi.⁵⁸ Apabila pengecoh tidak dipilih oleh peserta, maka dianggap tidak efektif dan perlu direvisi atau diganti. Ketiga indikator ini *point biserial*, *prop. correct*,

⁵⁷ Arikunto, *Loc. Cit*, 2020.

⁵⁸ *Ibid.*

dan *prop. endorsing* membentuk satu kesatuan yang saling melengkapi dalam mengevaluasi kualitas soal. Oleh karena itu, analisis berbasis ITEMAN sangat esensial untuk meningkatkan validitas, reliabilitas, dan daya guna instrumen evaluasi dalam proses pembelajaran.

BAB V

KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil analisis kualitas butir soal UAS Kimia kelas XI di SMA Negeri 1 Benai tahun ajaran 2023/2024 dapat disimpulkan bahwa sebagian besar soal, yaitu 80%, memiliki validitas tinggi hingga sangat tinggi, sedangkan sisanya masih perlu direvisi atau bahkan dibuang. Reliabilitas soal menunjukkan nilai alpha sebesar 0,793 yang termasuk dalam kategori tinggi, sehingga dapat dikatakan bahwa instrumen soal tersebut cukup konsisten. Dari segi tingkat kesukaran, mayoritas soal berada pada kategori sedang (56%), sementara 36% soal tergolong sukar dan 8% tergolong mudah. Ditinjau dari daya pembeda, sebanyak 64% soal mampu membedakan siswa dengan baik, sedangkan sisanya masih berada pada kategori cukup dan kurang. Selain itu, sebagian besar pengecoh dinyatakan efektif, meskipun masih terdapat beberapa opsi jawaban yang tidak efektif dan perlu dilakukan perbaikan agar soal lebih berkualitas.

B. Saran

Berdasarkan hasil penelitian ini, penulis menyarankan agar guru lebih teliti dalam menyusun soal ujian dengan memperhatikan tingkat kesukaran dan daya pembeda sehingga hasil evaluasi pembelajaran dapat lebih akurat dan adil. Pihak sekolah sebaiknya juga melakukan evaluasi secara berkala terhadap kualitas soal ujian agar instrumen yang digunakan benar-benar sesuai dengan standar pendidikan yang berlaku. Bagi siswa,

soal ujian yang berkualitas diharapkan dapat dimanfaatkan sebagai sarana untuk mengukur kemampuan diri sekaligus meningkatkan motivasi belajar. Bagi peneliti, penelitian ini menjadi pengalaman berharga dalam menguasai teknik analisis butir soal serta memperluas wawasan di bidang evaluasi pendidikan. Sementara itu, bagi peneliti selanjutnya, penelitian ini dapat dijadikan acuan untuk mengembangkan kajian lebih lanjut dengan memperluas objek, metode, maupun variabel penelitian sehingga hasil yang diperoleh dapat lebih komprehensif dan bermanfaat.

DAFTAR KEPUSTAKAAN

- Akhiruddin. (2019). *Belajar dan pembelajaran*. Kabupaten Gowa: Cahaya Bintang Cemerlang.
- Alti, Rizka Putri, Zulyuri, dan Violita. (2022). —Analisis kualitas butir soal ujian tengah semester mata pelajaran biologi kelas X di MAN 1 Solok Selatan.‖ *Jurnal Metaedukasi* 4, no. 2.
- Amalia, F, dan T Nugroho. (2022). —Evaluasi daya pembeda soal dalam pembelajaran kimia.‖ *Jurnal Pendidikan dan Pembelajaran Kimia* 10, no. 2
- Amelia, R, Anggi Ristiyana, dan Sri Rejeki Astuti. (2021). —Chemistry learning outcomes assessment: How is the quality of the tests made by the teacher?‖ *Journal of Educational Chemistry (JEC)* 3, no. 1
- Anggraeni, T. (2019). —Analisis tingkat kesukaran soal ujian akhir pada mata pelajaran IPA.‖ *Jurnal Penelitian dan Evaluasi Pendidikan* 7, no. 3: 197–205.
- Arbiatin, Eni, dan Mulabbiyah. (2020). —Analisis kelayakan butir soal tes penilaian akhir semester mata pelajaran matematika kelas VI di SDN 19 Ampenan tahun pelajaran 2019/2020.‖ *El Midad* 12, no. 2: 146–71.
- Arikunto, Suharsimi. (2018). *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.
- . (2019). *Dasar-dasar evaluasi pendidikan*. Revisi. Jakarta: Bumi Aksara.
- . (2020). *Prosedur penelitian suatu pendekatan praktik*. Jakarta: Bumi Aksara.

- Azwar, S. (2023). *Reliabilitas dan validitas instrumen penelitian (Edisi Revisi)*. Yogyakarta: Pustaka Pelajar.
- Bano, V O, D N Marambaawang, dan Y Njoeroemana. (2022). —Analisis kriteria butir soal ujian sekolah mata pelajaran IPA di SMP Negeri 1 Waingapu.‖ *Ideas: Jurnal Pendidikan, Sosial, dan Budaya* 8, no. 1: 145.
- Chasanah, Nur, Wahono Widodo, dan Nadi Suprpto. (2022). —Pengembangan instrumen asesmen literasi sains untuk mendeskripsikan profil peserta didik.‖ *PENDIPA: Journal of Science Education* 6, no. 2: 474–83. <https://doi.org/10.33369/pendipa.6.2.474-483>.
- Fauziana, A, dan Andhita Dessy Wulansari. (2021). —Analisis kualitas butir soal ulangan harian di sekolah dasar dengan model Rasch.‖ *Jurnal Kependidikan Dasar Islam Berbasis Sains* 6, no. 1.
- Fiska, J M, Y Hidayati, N Qomaria, dan W P Hadi. (2021). —Analisis butir soal ulangan harian IPA menggunakan software Anates pada pendekatan teori tes klasik.‖ *Natural Science Education Research* 4, no. 1: 65–76.
- Gusti, U A. —Analisis kualitas soal UAS mata pelajaran IPA di SMP Islam Khaira Ummah Kota Padang kelas VIII tahun ajaran 2019/2020.‖ *Biopedagogia* 3, no. 2 (2021): 81–87.
- Hidayat, F. (2022). —Analisis kualitas butir soal kimia untuk ujian akhir semester di SMA Negeri 3 Medan.‖ *Jurnal Evaluasi dan Penilaian Pendidikan* 10, no. 2: 100–110.
- Ikawati, H D, A Jayadi, dan Hermians. (2022). —Analisis kualitas tes dan butir soal sejarah di SMAN 1 Praya Timur.‖ *Jurnal Pendidikan dan Konseling* 4: 1349–58.

- Kaka, Leonardus, Oktoviana Bano, dan Yohana Njoeroemana. (2024). —Efektivitas analisis butir soal pilihan ganda menggunakan aplikasi anates di SMPN 2 Kanatang.‖ *Jurnal Inovasi Penelitian* 4, no. 9.
- Kemendikbud RI. (2014). *Kimia untuk SMA/MA kelas XI*. Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia.
- Magdalena, I, D Oktavia, dan P Nurjamilah. (2021). —Analisis evaluasi sumatif dalam pembelajaran tematik siswa kelas VI SDN Batujaya di era pandemi Covid-19.‖ *Arzusin* 1, no. 1.
- Mustaqim, dan Hidayu Sulisti. (2024). —Analisis butir soal PAS matematika peminatan: Daya pembeda, tingkat kesukaran, dan kualitas pengecoh.‖ *Al-Adad: Jurnal Tadris Matematika* 3, no. 1.
- Pahmi, S, C Hudaya, dan A Jaya. (2023). —Pengaruh model pembelajaran Tefa (Teaching Factory) dalam pembelajaran fisika untuk meningkatkan hasil belajar siswa SMK.‖ *Jurnal Hasil Kajian, Inovasi, dan Aplikasi Pendidikan Fisika* 9, no. 1: 1–15.
- Paskalin, G, dan M M I Susanti. (2020). —Analisis butir soal materi gaya kelas IV di SD Negeri Kentungan.‖ *Didaktika: Jurnal Pendidikan Sekolah Dasar* 3, no. 1: 23–32.
- Purniasari, L, M Masykuri, D Sri, dan R D Ariani. (2021). —Analisis butir soal ujian sekolah mata pelajaran kimia SMA N 1 Kutowinangun tahun pelajaran 2019/2020 menggunakan model Iteman dan Rasch.‖ *Jurnal Pendidikan Kimia* 10, no. 2: 205–14.
- Rahmawati, D, dan N Yuliani. (2020). —Peningkatan reliabilitas butir soal melalui analisis empiris.‖ *Jurnal Evaluasi Pendidikan* 12, no. 2: 110–20.

Republik Indonesia. (2020). *Undang-Undang Republik Indonesia Nomor 23 Tahun 2020 tentang Sistem Pendidikan Nasional*. Jakarta: Pemerintah Pusat.

Sabara, A, F M Yusuf, J Ahmad, E Nusantara, dan C Lamangajto. (2022). —Validitas perangkat pembelajaran berbasis STEAM materi struktur dan fungsi jaringan tumbuhan untuk meningkatkan penguasaan konsep.‖ *Jambura Edu Biosfer Journal* 4, no. 2: 76–85.

Santoso, A, dan M Dewi. (2023). —Efektivitas pengecoh pada soal ujian pilihan ganda.‖ *Jurnal Ilmu Pendidikan* 21, no. 3: 30–40.

Santoso, Agung. (2023). —Rumus Slovin: Panacea masalah ukuran sampel?‖ *Suksma: Jurnal Psikologi Universitas Sanata* 4, no. 2.

Sappaile, B I, dan T Pristiwaluyo. (2019). —Analisis butir soal ujian sekolah berstandar nasional dengan pendekatan klasik dan teori respon butir mata pelajaran matematika.‖ In *Proceedings of National Seminar Research and Community Service Institue Universitas Negeri Makassar*, 15:246–53.

Sholiha, Rohmatus, dan Maulfi Syaiful Rizal. —Pelaksanaan dan hambatan evaluasi formatif dalam pembelajaran menulis teks laporan hasil observasi di SMK PGRI 3 Malang.‖ *Jurnal Pendidikan Bahasa* 12, no. 1 (1 Juli 2023): 192–209. <https://doi.org/10.31571/bahasa.v12i1.5719>.

Simamora, Hendrik, Hartono, dan Effendi. (2021). —Analisis kualitas butir soal buatan guru kimia pada tes ujian tengah semester ganjil kelas XII MIPA.‖ *Hydrogen: Jurnal Kependidikan Kimia* 9, no. 1.

Siregar, Nurul Hidayah, Remiswal, dan Khadijah. (2024). —Analisis butir soal ujian tengah semester pada mata pelajaran pendidikan agama Islam.‖ *Urwatul Wutsqo: Jurnal Studi Kependidikan dan Keislaman* 13, no. 2.

Sridadi, Riky Dwihandaka, dan Ariyo Bagiastomo. (2020). —Evaluasi tes hasil belajar ulangan akhir semester genap mata pelajaran PJOK kelas VIII SMP N 1 Ngemplak tahun ajaran 2017/2018 dengan analisis butir soal.‖ *Jurnal Pendidikan Jasmani Indonesia* 16, no. 1: 28–40. <https://doi.org/10.21831/jpji.v16i1.31416>.

Sudarsono, S. (2022). —Analisis kualitas prediksi soal ujian nasional hasil diklat teknis substantif UN guru matematika MTs ditinjau dari tingkat kesukaran, daya pembeda dan kualitas pengecoh.‖ *Widyadewata* 4, no. 1: 47–54.

Sugiyono. (2017). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.

Sukardi. (2016). *Metodologi penelitian pendidikan*. Jakarta: Bumi Aksara.

Taufik, A N, L Berlian, A Fathia, dan E Y Rasyidin. (2022). —Pengembangan instrument tes literasi sains pada tema Yuk Makan Talas Beneng untuk calon guru IPA.‖ *Jurnal Aksioma* 11, no. 2: 141–50.

Wibowo, A. (2019). —Analisis validitas butir soal ujian sekolah pada mata pelajaran kimia.‖ *Jurnal Pendidikan Sains* 7, no. 1: 25–35.

Widoyoko, S E P. (2018). *Penilaian hasil pembelajaran di sekolah*. Jakarta: Pustaka Pelajar.

Wulandari, S. (2022). —Pengaruh daya pembeda terhadap kualitas soal pilihan ganda di sekolah menengah.‖ *Jurnal Penilaian Pendidikan* 10, no. 1: 78–86.