



Analisis Kualitas Instrumen Evaluasi Hasil Belajar Guru Kimia Kelas X SMAN 4 Enrekang

Quality Analysis of Class X Chemistry Teacher Learning Outcomes Evaluation Instrument at SMAN 4 Enrekang

Nurmia^{1*}, Maryono², Muhammad Anwar³,

^{1,2,3} Jurusan Kimia, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar

*Email: nurmia25@unm.ac.id

ABSTRAK

Penelitian ini adalah penelitian deskriptif yang bertujuan untuk mengetahui kualitas instrumen evaluasi hasil belajar siswa berdasarkan validitas, realibilitas, daya pembeda, indeks kesukaran, dan efektifitas pengecoh, serta menentukan presentase soal yang termasuk *High order thinking skill* di kelas X MIPA SMA Negeri 4 Enrekang. Subjek penelitian ini adalah seluruh peserta didik kelas X MIPA SMAN 4 Enrekang yang berjumlah 123 peserta didik. Pengambilan data dilakukan dengan menggunakan tes hasil belajar dan pedoman wawancara guru. Data tersebut dianalisis menggunakan analisis deskriptif. Hasil penelitian menunjukkan bahwa kualitas instrumen evaluasi hasil belajar guru kimia memiliki validitas isi yang tinggi. Adapun kelayakan tes hasil belajar kimia berdasarkan: (1) Jumlah kategori item soal validitas tinggi adalah 10%; (2) Reliabilitas item soal memiliki kategori tinggi dengan $r_{11} = 0,78$; (3) Jumlah item soal kategori daya pembeda baik sekali adalah 6,67%; (4) Jumlah item soal kategori tingkat kesukaran sangat baik adalah 24,17%; (5) Untuk efektifitas fungsi opsi kunci efektif 46,67%, efektifitas fungsi opsi kunci tidak efektif 53,33%, dan efektifitas fungsi opsi pengecoh efektif 50,83%, efektifitas fungsi opsi tidak efektif 49,17%, (6) Soal-soal yang termasuk kedalam soal-soal *High Order Thinking Skill* sebesar 23,33%.

Kata Kunci : Analisis 1, Validitas 2, Realibilitas 3, Tingkat Kesukaran 4, Daya Pembeda 5, Efektifitas Pengecoh 6, High Order Thinking Skill 7.

ABSTRACT

This research is a descriptive study that aims to determine the quality of the evaluation instrument for student learning outcomes based on validity, reliability, discriminating power, difficulty index, and the effectiveness of the deluder, as well as determining the percentage of questions that include High order thinking skills in class X MIPA SMA Negeri 4 Enrekang. The subjects of this study were all 123 students in class X MIPA at SMAN 4 Enrekang. Data collection was carried out using learning achievement tests and teacher interview guidelines. The data were analyzed using descriptive analysis. The results showed that the quality of the chemistry teacher's learning outcomes evaluation instrument had high content validity. The feasibility of the chemistry learning outcomes test is based on: (1) The number of item categories with high validity is 10%; (2) The reliability of the question item has a high category with $r_{11} = 0.78$; (3) The number of items in the category of excellent discrimination is 6.67%; (4) The number of items in the category of very good difficulty level is 24.17%; (5) The effectiveness of the effective lock option function is 46.67%, the effectiveness of the ineffective option function is 53.33%, and the effectiveness of the effective deluder option function is 50.83%, the effectiveness of the ineffective option function is 49.17%, (6) Problem -questions that are included in the High Order Thinking Skill questions are 23.33%.

Keywords : Analysis 1, Validity 2, Reliability 3, Difficulty Level 3, Distinguishing Power 4, Detractor Effectiveness 5, High Order Thinking Skill 6.

PENDAHULUAN

Secara harfiah kata evaluasi berasal dari bahasa Inggris *evaluation*; dalam bahasa Arab; *al-taqdir*; dalam bahasa Indonesia berarti; penilaian. Akar katanya adalah *value*; dalam bahasa Arab; *al-qimah*; dalam bahasa Indonesia berarti; nilai (Wulan dan Rusdiana, 2014). analisis kualitas soal adalah sebuah proses yang harus ditempuh dan dilaksanakan oleh seorang guru untuk mengetahui sejauh mana kualitas butir soal yang diberikan (Farida dan Musyarofah, 2021).

Kualitas hasil evaluasi dalam prakteknya dipengaruhi oleh beberapa hal, antara lain tes yang digunakan. Menurut Arikunto (2013), tes merupakan alat atau prosedur yang digunakan untuk mengetahui atau mengukur sesuatu dalam suasana, dengan cara dan aturan-aturan yang sudah ditentukan. Peranana tes dalam rangka peningkatan mutu pendidikan sangat penting karena tes adalah suatu prosedur yang sistematis untuk mengamati perilaku seseorang dan dijelaskan dalam bentuk skor atau sistem kategori. Suatu tes yang dapat dikatakan baik sebagai alat pengukur, harus memenuhi persyaratan tes, yaitu memiliki validitas, realibilitas, objektivitas, praktibilitas, dan ekonomis (Arikunto, 2013).

Tes sebagaimana diuraikan di atas, adalah tes yang proses pembuatannya memerlukan waktu, direncanakan secara khusus, dilaksanakan secara prosedur yang benar, dibuat oleh tenaga ahli (guru-guru) yang menguasai materi atau substansi dari pelajaran yang akan dibuat tesnya, dan menguasai teknik dalam pembuatan tes yang baik dan benar. Dilihat dari cara penyusunannya, tes dapat dibedakan atas dua jenis, yaitu tes buatan guru (*teacher-made test*) dan tes standar (*standardized test*) (Arikunto, 2013). Tes juga merupakan salah satu cara untuk mengukur kompetensi

seseorang setelah mempelajari sesuatu (Utomo, 2018).

Penilaian yang dilaksanakan pada akhir semester genap oleh satuan pendidikan untuk mengukur pencapaian kompetensi peserta didik pada akhir semester adalah Penilaian Akhir Tahun (PAT). Cakupan materi pada penilaian akhir tahun meliputi seluruh indikator yang merepresentasikan kompetensi dasar pada semester genap pada tingkat kelas yang sama (Kemdikbud, 2017).

Tujuan penilaian ini adalah untuk mengetahui hasil kognitif yang dicapai oleh para peserta didik, yakni seberapa jauh peserta didik telah mencapai kompetensi yang ditetapkan. Proses penilaian aspek kognitif dilakukan dengan teknik ujian yang menggunakan instrumen soal. Instrumen soal yang digunakan oleh guru umumnya masih menggunakan butir soal yang ada di buku teks, lembar kegiatan peserta didik, atau kumpulan soal yang telah diberikan peserta didik. Kondisi ini membuat peserta didik cenderung hanya menggunakan ingatan dalam menyelesaikan soal (Hartini & Sukardjo, 2015). Penilaian pengetahuan merupakan penilaian untuk mengukur kemampuan peserta didik berupa pengetahuan faktual, konseptual, prosedural, dan metakognitif, serta kecakapan berpikir tingkat rendah sampai tinggi. Penilaian ini berkaitan dengan ketercapaian KD pada KI-3 yang dilakukan oleh guru mata pelajaran (Kemendikbud, 2017).

Salah satu cara agar peserta didik mampu berpikir tingkat tinggi adalah dengan menyajikan soal-soal yang memiliki kriteria *High order thinking skills*. *High order thinking skills* (HOTS) atau keterampilan

berpikir tingkat tinggi merupakan bagian dari taksonomi Bloom yang termasuk dalam kategori *analyze* (C4), *evaluate* (C5) dan *create* (C6) yang dapat digunakan dalam penyusunan soal. Pendidik harus memiliki pengetahuan dan keahlian untuk menunjang pekerjaannya, sehingga dapat mengembangkan keterampilan berpikir tingkat tinggi peserta didik tingkat SMA (Aydin dan Yilmaz, 2010). Terkadang tidak semua kompetensi dasar dapat dibuatkan soal yang bersifat *High Order Thinking Skills*. Oleh karena itu, kejelian dan ketelitian dari tenaga pendidik sangat diperlukan.

Berdasarkan dokumentasi yang dilakukan di SMAN 4 Enrekang terdapat empat kelas yaitu X MIPA 1, X MIPA 2, X MIPA 3, dan X MIPA 4. Guru mata pelajaran kimia membuat sendiri instrumen penilaian hasil belajar untuk ulangan harian, ulangan tengah semester, dan ulangan akhir semester. Bentuk soal yang digunakan di sekolah dalam Ujian Akhir Semester (UAS) adalah bentuk tes objektif (pilihan ganda). Soal dalam Ujian Akhir Semester harus memiliki kualitas yang baik agar dapat mengukur kemampuan hasil belajar peserta didik secara tepat dan akurat. Untuk itu, soal harus dianalisis untuk mengetahui kualitas soal tersebut. Menurut Arikunto (2013), analisis soal bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang baik, dan soal yang jelek. Dengan analisis soal dapat diperoleh informasi tentang kekurangan sebuah soal dan petunjuk untuk mengadakan perbaikan.

Berdasarkan uraian diatas, maka telah dilakukan suatu penelitian yang berjudul tentang Analisis Kualitas Instrumen Evaluasi Hasil Belajar Guru Kimia SMAN 4 Enrekang.

METODE

Penelitian ini merupakan penelitian deskriptif. Penelitian deskriptif merupakan

penelitian yang ditujukan untuk menggambarkan fenomena yang ada secara verbal atau dengan kalimat dan numerik, yang berlangsung pada saat ini atau masa lampau.

Penelitian dilakukan di SMAN 4 Enrekang pada semester genap Tahun ajaran 2021/2022. Instrumen penelitian yang digunakan pada penelitian ini adalah lembar validasi isi dan pedoman wawancara. Wawancara dilakukan dengan guru kimia kelas X MIPA SMAN 4 Enrekang untuk mengetahui bagaimana guru kimia membuat instrumen evaluasi hasil belajar yang diberikan kepada peserta didik.

Teknik yang digunakan peneliti dalam pengumpulan data adalah dokumentasi dan wawancara. Teknik dokumentasi merupakan suatu teknik pengumpulan data dengan menghimpun dan menganalisis dokumen-dokumen, baik dokumen tertulis, gambar maupun elektronik. Teknik dokumentasi digunakan untuk mendapatkan data berupa lembar Soal Ujian Akhir Semester Genap mata pelajaran Kimia kelas X MIPA SMAN 4 ENREKANG dan lembar jawaban peserta didik. Kemudian dilakukan wawancara guru untuk mengetahui proses pembuatan instrumen evaluasi hasil belajar guru kimia kelas X SMAN 4 Enrekang. Wawancara merupakan teknik pengumpulan data sebagai studi pendahuluan untuk mendapatkan permasalahan yang dialami.

HASIL DAN PEMBAHASAN

Hasil penelitian menunjukkan bahwa kualitas hasil belajar guru kimia memiliki kualitas yang tinggi. Kualitas

intrumen tersebut dapat dilihat dari hasil analisis validitas, reliabilitas, daya pembeda, indeks kesukaran, dan efektifitas pengecoh serta pengelompokan soal berdasarkan soal-soal *High Order Thinking Skill*.

Terdapat beberapa faktor yang menyebabkan butir soal menjadi tidak valid,

ada tiga faktor yang mempengaruhi validitas pada hasil tes yaitu faktor instrumen yang dipakai untuk tes, faktor administrasi dan penskoran, serta faktor dari jawaban peserta didik (Prawiki dan Helendra, 2022).

Tabel 1. Uji Validitas

Tingkat	Jumlah	Nomor Soal	Persentase (%)
Tinggi	3	10, 24, 25	10
Cukup	17	3, 5, 7, 8, 9, 11, 12, 13, 16, 17, 18, 19, 20, 23, 26, 29, 30	56,67
Rendah	16	1, 2, 4, 14, 22, 27	20
Sangat rendah	4	6, 15, 21, 28	13,33
Jumlah	30	30	100
Rata-rata		0,43	

Validitas soal berkaitan dengan sejauh mana butir soal dapat mengukur hasil belajar yang hendak diukur (Tarmizi, dkk, 2020).

dari siswa. Untuk uji validitas pada penelitian ini menggunakan rumus korelasi *product moment*. Jika nilai r_{pbis} lebih tinggi dari nilai r tabel, sehingga soal tersebut termasuk dalam kriteria valid (Purniasari, dkk,

2021). Diperoleh hasil r hitung $>$ rtabel (0.176) maka dinyatakan valid, begitu juga arti sebaliknya. Butir soal dikatakan memiliki kualitas yang baik, apabila soal tersebut memiliki validitas tes termasuk kategori valid (Fauziana dan Wulansari, 2021) Peserta didik kelas X MIPA SMAN 4 Enrekang berjumlah 123 orang.

Tabel 2. Reliabilitas

Kelas	Reliabilitas	Jumlah Butir Soal
X MIPA	0,78	30

Reliabilitas merupakan ketetapan atau keajegan alat tersebut dalam penilaian apa yang dinilainya. Artinya, kapan pun alat penilaian tersebut digunakan akan memberikan hasil yang relative sama. Reliabilitas berhubungan dengan keandalan dan keajegan hasil. Butir soal akan memiliki tingkat keandalan yang tinggi, apabila tes memberikan hasil yang korelatif tetap (ajeg) (Fatayah, dkk, 2022).

Berdasarkan hasil analisis reliabilitas pada soal-soal pilihan ganda UAS Kimia tahun ajaran 2021-2022 kelas X MIPA SMAN 4 Enrekang dapat digolongkan reliabilitasnya tinggi dengan besar koefisien korelasi $r_{11} = 0,78$. Semakin tinggi indeks reliabilitas suatu tes, semakin tinggi juga keajegan atau ketepatannya. Apabila realibilitas pada suatu soal menunjukkan angka diantara $0,70 < r < 0,89$ maka korelasi soal memiliki

realibilitas yang tinggi. Azwar (1996) mengartikan realibilitas sebagai keterpercayaan, keterandalan, keajegan, konsistensi, dan kestabilan. Maka dapat diketahui bahwa instrument evaluasi hasil belajar guru

kimia kelas X SMAN 4 Enrekang tersusun oleh butir-butir soal yang mampu menunjukkan keajegan hasil atas pengukuran yang dilakukan sehingga soal dapat dipercaya.

Tabel 3. Daya Pembeda

Daya Pembeda	Kriteria	Jumlah	Nomor Soal	Persentase (%)
0,00 – 0,200	Jelek	7	6, 11, 21, 23, 28, 29, 30	23,33
0,21 – 0,40	Cukup	10	3, 4, 7, 8, 14, 15, 17, 19, 22, 24	33,33
0,41 – 0,70	Baik	11	1, 2, 5, 9, 10, 12, 16, 18, 25, 26, 27	36,67
0,71 – 1,00	Baik Sekali	2	13, 20	6,67
Jumlah		30	30	100
Rata-rata		0,36		

Berdasarkan dari hasil penelitian yang diperoleh pada Tabel 3 dilihat dari daya pembeda soal UAS kelas X MIPA sebanyak 11 soal sudah termasuk baik. Artinya peserta didik kelas bawah lebih banyak menjawab pilihan jawaban yang benar dibandingkan dengan peserta didik kelas atas. Daya beda bertujuan untuk mengetahui kemampuan setiap peserta didik, dan validitas sebagai alat ukur untuk mengetahui seberapa berfungsi butir soal tersebut dalam membedakan peserta didik yang memiliki kemampuan tinggi dan

kemampuan rendah (Nurhalimah, dkk, 2022). Nilai daya pembeda dinyatakan melalui indeks daya pembeda. Makin tinggi atau makin besar indeks daya pembeda soal, makin besar soal tersebut dapat membedakan antara kelas atas dan kelas bawah. Karena intrumen evaluasi hasil belajar kimia didominasi oleh daya pembeda dengan kriteria cukup, baik dan baik sekali maka daya pembeda mampu membedakan antara kelas atas dan kelas bawah.

Tabel 4. Indeks Kesukaran

Indeks Kesukaran	Jumlah	Nomor Soal	Persentase (%)
0,00 – 0,30 adalah soal sukar	3	6, 14, 28	10
0,31 -0,70 adalah soal sedang	10	1, 2, 9, 12, 13, 15, 18, 20, 26, 27	33,33
0,71 – 1,00 adalah soal mudah	17	3, 4, 5, 7, 8, 10, 11, 16, 17, 19, 21, 22, 23, 24, 25, 29, 30	56,67

Indeks Kesukaran	Jumlah	Nomor Soal	Persentase (%)
Jumlah	30	30	100
Rata-rata		0,68	

Menurut Arifin (2012) untuk memperoleh prestasi belajar yang baik, sebaiknya proporsi antara tingkat kesukaran soal tersebar secara normal. Proporsi soal diatur untuk soal sukar 25%, soal sedang 50%, soal mudah 25% atau soal sukar 20%, soal sedang 60%, soal mudah 20%, atau soal sukar 15 %, soal sedang 70%, soal mudah 15%. Dalam suatu evaluasi pembelajaran dibutuhkan juga evaluasi proses penilaian dan pengukuran siswa. Analisis kesukaran soal sangat dibutuhkan sekali. Karena untuk mengetahui apakah peserta didik bisa mengerjakan soal yang diberikan atau justru soal yang diberikan terlalu mudah bahkan terlalu sukar (Fatimah dan Alfath, 2019). Sebuah paket soal harus memiliki proporsi seimbang antara kategori sukar, sedang, dan mudah (Anita, dkk, 2018).

Berdasarkan Tabel 4 pada kelas X MIPA diketahui sebanyak 3 soal

kategori sukar, 10 soal kategori sedang, dan 17 soal kategori mudah. Hasil perbandingan tersebut menunjukkan instrument evaluasi hasil belajar UAS kelas X MIPA mempunyai proporsi yang tidak seimbang dengan kata lain dapat dikatakan soal UAS pada kelas X MIPA didominasi oleh tingkat kesukaran dengan kriteria mudah. Semakin rendah angka indeks kesukaran soal, maka soal tersebut semakin sukar karena sedikit peserta tes yang menjawab dengan benar soal tersebut. Soal yang baik adalah soal yang tidak terlalu sukar dan tidak terlalu mudah, soal yang terlalu mudah tidak merangsang peserta didik untuk mempertinggi usahanya dalam memecahkan soal. Sebaliknya, soal yang terlalu sukar akan menyebabkan peserta didik menjadi tidak memiliki semangat untuk mencobanya karena diluar kemampuannya.

Tabel 5. Efektifitas Pengecoh

Kategori IP	Jumlah Pengecoh	Persentase (%)
Sangat Baik = 76% - 125%	29	24,17
Baik = 51% - 75% atau 126% - 150%	20	16,67
Kurang Baik = 26% - 50% atau 151% - 175%	28	23,33
Jelek = 0% - 25% atau 176% - 200%	31	25,83
Sangat Jelek = Lebih dari 200%	12	10
Jumlah	30	100

Suatu distraktor (pengecoh) dapat dikatakan berfungsi baik jika jumlah pemilih kelompok atas dan kelompok bawah tidak kurang dari 2 orang peserta didik dan jumlah pemilih kelompok bawah harus lebih besar dari kelompok atas. Berdasarkan hasil analisis penelitian menunjukkan instrumen evaluasi hasil belajar memiliki efektifitas fungsi opsi pengecoh yang berfungsi efektif sebesar 50,83% dan pengecoh yang tidak efektif 49,17%. Hasil tersebut menunjukkan bahwa pengecoh pada instrumen evaluasi hasil

belajar kimia pada UAS X MIPA SMAN 4 Enrekang sebagian besar memiliki pengecoh yang efektif. Apabila terlalu banyak peserta didik yang memilih kunci jawaban atau pilihan jawaban tersebut maka pengecoh yang telah dibuat belum dapat dikatakan efektif (Simamora, dkk, 2021). Butir soal yang pengecohnya berfungsi sangat baik dan baik harus dipertahankan dengan memasukkannya kedalam bank soal (Muluki, dkk, 2020).

Tabel. 6 Pengelompokan Soal-soal *High Order Thinking Skill*

Nomor Soal	Level Kognitif
9, 12, 13, 14, 15, 24, 27	C4
Jumlah	7

Menurut Helmawati (2019), persentasi penilaian pembelajaran ada aspek kognitif hendaknya didistribusikan

sebagai berikut ; 1) C1-pengetahuan sebanyak 5%, 2) C2-pemahaman sebanyak 10%, 3) C3-penerapan sebanyak 45%, C4-analisis sebanyak 10%, 5) C5-evaluasi sebanyak 10%,

dan 6) C6-kreasi sebanyak 5%. Hasil penelitian tidak menunjukkan proporsi yang sama untuk persentasi penilaian pembelajaran. Hasil penelitian menunjukkan bahwa tidak ada item soal yang masuk dalam level kognitif C5

dan C6. Hasil ini menunjukkan bahwa kemampuan guru kimia dalam membuat instrument evaluasi hasil belajar UAS masih didominasi oleh unsur *Low Order Thinking Skills*. Kemampuan berpikir tingkat tinggi peserta didik sangat dipengaruhi oleh pembelajaran. Untuk meningkatkan kemampuan *High Oreder Thinking Skill* pada peserta didik maka perlu diberikan latihan untuk mengerjakan soal-soal dengan tipe *High Oreder Thinking Skill*, tipe soal tersebut dapat diberikan pada ujian harian, ujian tengah semester, dan ujian akhir semester (Risidiana, dkk, 2022)

KESIMPULAN DAN SARAN

A. Kesimpulan

Rata-rata *r product moment* validitas item soal 0,43; reliabilitas item soal memiliki kategori tinggi dengan $r_{11} = 0,78$; Rata-rata daya pembeda item soal 0,36; Rata-rata tingkat kesukaran item soal 0,68; Jumlah item soal kategori IP sangat baik adalah 24,17%,; Soal-soal yang termasuk kedalam soal-soal *High Order Thinking Skill* sebesar 23,33%.

B. Saran

Guru perlu ada perbaikan aspek konstruksi dan bahasa pada tes hasil belajar peserta didik dan Sekolah perlu mengadakan pelatihan pembuatan soal berkategori *High Order Thinking Skill*.

DAFTAR PUSTAKA

- Anastasia, A dan Urbina, S. 1997. *Tes Psikologi Edisi Bahasa Indonesia*. Jakarta: PT.Prenhalindo.
- Anita, Sulis tyowati, Zuldafrial. 2018. Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas. *Jurnal Pendidikan*. Vol. 16. No. 1.
- Arifin, Zainal. 2012. *Evaluasi Pembelajaran*. Jakarta: Direktorat Jenderal Pendidikan Islam Kementerian Agama
- Arikunto, Suharsimi. 2013. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Darma, Budi. 2021. *Statistika Penelitian Menggunakan SPSS (Uji Validitas, Uji Reliabilitas, Regresi Linier Sederhana, Regresi Linier Berganda, Uji t, Uji F, R²)*. Jakarta: GUEPEDIA.
- Daryanto. 2008. *Evaluasi Pendidikan*. Jakarta: PT Rineka Cipta.
- Farida, dan Anna Musyarofah. 2021. Validitas dan Reliabilitas dalam Analisis Butir Soal. *Jurnal Pendidikan Bahasa Arab*. Vol. No. 1.
- Fatayah, Ika Farida Yuliana, Layla Muf'idah. 2022. Analisis Validitas Dan Reliabilitas Dalam Mendukung Ketuntasan Belajar Model STEM. *Jurnal Buana Pendidikan*. Vol. 18. No. 1
- Fatimah, Laela Umi dan Alfaft, Khaeruddin. 2019. Analisis Kesukaran Soal, Daya Pembeda, dan Fungsi Distraktor. *Jurnal Komunikasi dan Pendidikan Islam*. Vol. 8. No. 2.
- Fauziana, Anis dan Wulansari, Anditha Dessy. 2021. Analisis Kualitas Butir Soal Ulangan Harian di Sekolah Dasar dengan Model Rasch. *Jurnal Kependidikan Dasar Islam Berbasis Sains*. Vol. 6. No. 1

- Kemdikbud. 2017. *Panduan Penilaian oleh Pendidik dan Satuan Pendidikan Sekolah Menengah Atas*. Jakarta: Direktorat Pembinaan SMA Ditjen Pendidikan Dasar dan Menengah.
- Kemendikbud. 2019. *Modul Penyusunan Soal HOTS*. Jakarta: Direktorat Jendral Pendidikan Dasar Menengah.
- Muluki, Ardilah, Patta Bundu, I Sukmawati. 2020. Analisis Kualitas Butir Tes Semester Ganjil Mata Pelajaran IPA Kelas IV Mi Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*. Vol. 4. No.1.
- Nurhalimah, Sri, Ynin Hidayati, Irsad Rosidi, Wiwin Puspita Hadi. 2022. Hubungan Antara Validitas Item dengan Daya Pembeda dan Tingkat Kesukaran Soal Pilihan Ganda PAS. *Jurnal Natural Science Educational Research*. Vol 4. No. 3.
- Prawiki, Suci Mitra dan Helendra. 2022. Analisis Kualitas Butir Soal Ujian Akhir Semester Ganjil Tahun Pelajaran 2020/2021 Mata Pelajaran Biologi Kelas X SMA Negeri 1. *Jurnal Biologi dan Pembelajarannya*. Vol. 17. No. 12.
- Purniasari, Laksmi, Muhammad Masykuri, Dan Sri Retno Dwi Ariani. 2021. Analisis Butir Soal Ujian Sekolah Mata Pelajaran Kimia SMAN 1 Kutowinangun Tahun Pelajaran 2019/2020 Menggunakan Model ITEMAN Dan RASCH. *Jurnal Pendidikan Kimia*. Vol. 10. No. 2.
- Risdiana, Aulia, Mariana Erna, Betty Holiwarni. 2022. Pengembangan Soal HOTS (Higher Order Thinking Skills) pada Materi Asam-Basa untuk Kelas XI SMA/MA Sederajat. *Jurnal Inovasi Pendidikan Kimia*. Vol. 2. No. 7.
- Retnawati, Heri. 2016. *Analisis Kuantitatif Instrumen Penilaian*. Yogyakarta: Parama Publishing.
- Sani, R. A. 2019. Cara Membuat Soal HOTS. Tangerang: Tira Smart
- Sartika dan Yusmaita. 2020. Pengembangan Assesmen Literasi Kimia Pada Materi Hukum-Hukum Dasar Kimia Dan Stoikiometri Kelas X Sma/Ma. *Jurnal Edukimia*. Vol. 2. No. 3.
- Simamora, Hendrik, Hartono, Effendi. 2021. Analisis Kualitas Butir Soal Buatan Guru Kimia Pada Tes Ujian Tengah Semester Ganjil Kelas XII MIPA. *Jurnal*

- Kependidikan Kimia*. Vo. 9. No.1
- Sudijono, Anas. 2011. *Evaluasi Pendidikan*. Jakarta: Raja Grafindo
- Sudirman. 1992. *Ilmu Pendidikan*. Bandung: PT Remaja Rosdakarya.
- Supriadi, Gito. 2020. *Pengembangan Instrumen Penilaian Berbasis Higher Order Thinking Skill (HOTS)*. Yogyakarta: Aswaja Pressindo.
- Surapranata, Sumarna. 2009. *Validasi, Reliabilitas, dan Interpretasi Hasil Tes*. Bandung: Remaja Rosdakarya.
- Surapranata, Sumarna. 2005. *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes*. Bandung: PT Remaja Rosdakarya.
- Surapranata, sumarna. 2005. *Panduan penulisan testertulis implementasi kurikulum 2004*. Bandung: remja rosakarya.
- Taramizi, Pebrian, Panut Setiono, Yuli Amaliah, Arief Agrian. 2020. Analisis Butir Soal Pilihan Ganda Tema Sehat Itu Penting Kelas V Sd Negeri 04 Kota Bengkulu. *Elementary School Education Journal*. Vol. 4. No. 2.
- Utomo, Budi. 2018. Analisis Validitas Isi Butir Soal sebagai Salah Satu Upaya Peningkatan Kualitas Pembelajaran di Madrasah Berbasis Nilai-Nilai Islam. *Jurnal Pendidikan Matematika*. Vol. 2. No. 1.
- Wulan, Ratna Elis Dan Rusdiana. 2014. *Evaluasi Pembelajaran Dengan Pendekatan Kurikulum 2013*. Bandung: Pustaka Setia Bandung.