

**UJI KUALITAS SOAL EVALUASI MATA PELAJARAN
BIOLOGI SMA YANG DIKEMBANGKAN MENGGUNAKAN
GEMINI AI DENGAN ANALISIS *RASCH MODEL***



**MUTIARA SALSABILA WARMAN
NIM. 21031025/2021**

**DEPARTEMEN BIOLOGI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2026**

**UJI KUALITAS SOAL EVALUASI MATA PELAJARAN
BIOLOGI SMA YANG DIKEMBANGKAN MENGGUNAKAN
GEMINI AI DENGAN ANALISIS *RASCH MODEL***

SKRIPSI

*Diajukan sebagai salah satu persyaratan untuk memperoleh gelar
Sarjana Pendidikan*



Oleh:

**MUTIARA SALSABILA WARMAN
NIM. 21031025/2021**

**PROGRAM STUDI PENDIDIKAN BIOLOGI
DEPARTEMEN BIOLOGI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2026**

HALAMAN PERSETUJUAN SKRIPSI

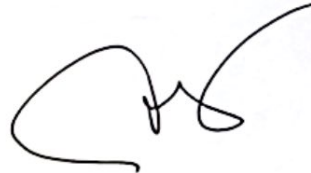
Judul : Uji Kualitas Soal Evaluasi Mata Pelajaran Biologi SMA Yang
Dikembangkan Menggunakan Gemini AI Dengan Analisis Rasch
Model
Nama : Mutiara Salsabila Warman
NIM/TM : 21031025/2021
Program Studi : Pendidikan Biologi
Departemen : Biologi
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Mengetahui,
Kepala Departemen Biologi FMIPA UNP



Dr. Dwi Hilda Putri, S.Si., M.Biomed
NIP. 197508152006042001

Padang, 23 April 2026
Disetujui oleh,
Dosen Pembimbing



Relsas Yogica, M.Pd
NIP. 199006022015041004

PENGESAHAN LULUS UJIAN SKRIPSI

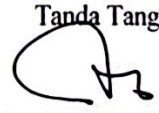
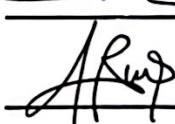
Nama : Mutiara Salsabila Warman
NIM : 21031025
Program Studi : Pendidikan Biologi
Departemen : Biologi
Fakultas : Matematika dan Ilmu Pengetahuan Alam

UJI KUALITAS SOAL EVALUASI MATA PELAJARAN BIOLOGI SMA YANG DIKEMBANGKAN MENGGUNAKAN GEMINI AI DENGAN ANALISIS *RASCH MODEL*

Dinyatakan telah lulus setelah dipertahankan di depan Tim Penguji Skripsi
Program Studi Pendidikan Biologi, Departemen Biologi
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Padang

Padang, 04 Mei 2026

Tim Penguji

	Nama	Tanda Tangan
Ketua	: Relsas Yogica, M.Pd	
Anggota	: Prof. Dr. H. Syamsurizal, M.Biomed	
Anggota	: Dr. Rahmawati D., S.Pd, M.Pd	

SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : Mutiara Salsabila Warman
NIM/TM : 21031025/2021
Program Studi : Pendidikan Biologi
Departemen : Biologi
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa, skripsi saya yang berjudul **“Uji Kualitas Soal Evaluasi Mata Pelajaran Biologi SMA yang Dikembangkan Menggunakan Gemini AI dengan Analisis *Rasch Model*”** adalah benar hasil karya sendiri, bukan hasil plagiat dari karya orang lain. Sepanjang pengetahuan saya tidak terdapat karya yang dituliskan dan diterbitkan orang lain kecuali sebagai acuan atau kutipan dengan mengikuti aturan penulisan karya ilmiah yang benar.

Demikian pernyataan ini saya buat dengan penuh kesadaran dan rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Padang, 06 Mei 2026

Diketahui oleh,
Kepala Departemen Biologi



Dr. Dwi Hilda Putri, S.Si., M.Biomed.
NIP. 197508152006042001

Saya yang menyatakan,



Mutiara Salsabila Warman
NIM. 21031025

ABSTRAK

Mutiara Salsabila Warman: Uji Kualitas Soal Evaluasi Mata Pelajaran Biologi SMA Yang Dikembangkan Menggunakan Gemini AI Dengan Analisis Rasch Model

Penelitian ini dilatarbelakangi oleh masih terbatasnya variasi soal evaluasi serta belum optimalnya jaminan kualitas soal yang dihasilkan oleh AI. Penelitian ini bertujuan untuk menganalisis kualitas soal evaluasi mata pelajaran Biologi tingkat SMA yang dikembangkan menggunakan teknologi kecerdasan buatan, yaitu Gemini AI, dengan pendekatan Model Rasch. Penelitian ini menggunakan metode deskriptif kuantitatif yang dilaksanakan di SMA Pembangunan Laboratorium UNP dengan subjek siswa kelas X, XI, dan XII.

Instrumen penelitian berupa soal pilihan ganda yang dikembangkan menggunakan Gemini AI melalui teknik prompt *RISE*, kemudian divalidasi oleh ahli dan diuji cobakan kepada peserta didik. Data dianalisis menggunakan pemodelan Rasch dengan indikator validitas, reliabilitas, indeks separasi, tingkat kesukaran soal, fungsi distraktor, serta analisis *Differential Item Functioning* (DIF).

Hasil penelitian menunjukkan bahwa secara umum soal yang dikembangkan memiliki kualitas yang baik. Sebagian besar butir soal memenuhi kriteria kesesuaian model (fit) berdasarkan nilai Outfit MNSQ, Outfit ZSTD, dan *Point Measure Correlation*. Nilai reliabilitas menunjukkan kategori tinggi, sedangkan indeks separasi mampu membedakan kemampuan siswa dengan baik. Tingkat kesukaran soal bervariasi dari mudah hingga sulit. Hasil analisis DIF menunjukkan adanya beberapa butir soal yang terindikasi bias. Selain itu, masih terdapat beberapa butir soal dengan distraktor yang belum berfungsi optimal, sehingga diperlukan revisi agar instrumen lebih efektif dan objektif.

Kata Kunci: biologi SMA, evaluasi pembelajaran, Gemini AI, kualitas soal, Model Rasch.

ABSTRACT

Mutiara Salsabila Warman: *Analysis of the Quality of High School Biology Test Items Developed Using Gemini AI Based on the Rasch Model*

This study was motivated by the limited variation of evaluation questions and the lack of optimal quality assurance of items generated by artificial intelligence. This study aims to analyze the quality of high school Biology assessment items developed using artificial intelligence technology, namely Gemini AI, with the Rasch Model approach. This research employed a quantitative descriptive method conducted at SMA Pembangunan Laboratorium UNP, with participants consisting of students from grades X, XI, and XII.

The research instrument consisted of multiple-choice questions developed using Gemini AI through the RISE prompt technique, which were then validated by experts and tested on students. The data were analyzed using Rasch modeling based on indicators of validity, reliability, separation index, item difficulty level, distractor effectiveness, and Differential Item Functioning (DIF).

The results indicate that, in general, the developed items have good quality. Most items meet the model fit criteria based on Outfit MNSQ, Outfit ZSTD, and Point Measure Correlation values. The reliability value is categorized as high, indicating consistent measurement, while the separation index shows the instrument's ability to distinguish students' ability levels effectively. The item difficulty levels vary from easy to difficult. The DIF analysis indicates that several items are biased. In addition, some items still have non-functioning distractors, indicating the need for revisions to improve the effectiveness and objectivity of the instrument.

Keywords: *biology education, Gemini AI, item quality, learning evaluation, Rasch model.*

KATA PENGANTAR

Puji dan syukur peneliti ucapkan atas kehadiran Allah SWT, berkat limpahan rahmat, hidayah, serta karunia-Nya yang telah memberikan kekuatan dan kemudahan bagi peneliti sehingga dapat menyelesaikan skripsi dengan judul "Uji Kualitas Soal Evaluasi Mata Pelajaran Biologi SMA yang Dikembangkan Menggunakan Gemini AI dengan Analisis *Rasch Model*".

Peneliti juga mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah berkontribusi dalam penyelesaian skripsi ini, baik melalui bimbingan, saran, masukan, ide, maupun dorongan semangat yang sangat berarti. Ucapan terima kasih ini secara khusus peneliti tujukan kepada pihak-pihak berikut.

1. Bapak Relsas Yogica, M.Pd., selaku dosen pembimbing skripsi yang telah memberikan saran dan masukan kepada peneliti serta kesediaan meluangkan waktu untuk membimbing dan memberi motivasi kepada peneliti.
2. Bapak Prof. Dr. H. Syamsurizal, M. Biomed dan Ibu Dr. Rahmawati D., S.Pd, M.Pd selaku dosen penguji dan validator yang telah memberikan kritik dan saran yang sangat membangun bagi penulis.
3. Ibu Dr. Helendra, M.S., selaku dosen penasihat akademik yang telah memberikan arahan selama kegiatan perkuliahan berlangsung dan memberi motivasi kepada peneliti.
4. Pimpinan, Staf Pengajar, Karyawan, serta Laboran Departemen Biologi FMIPA UNP yang telah memberikan kemudahan dalam penyusunan skripsi.

5. Kepala Sekolah, Wakil Kepala Sekolah, Majelis Guru serta peserta didik SMA Pembangunan Laboratorium UNP yang telah memberikan kemudahan dalam penyusunan skripsi.
6. Orang tua penulis, Papa Hendri Warman dan Mama Nanik, terima kasih atas doa yang tak pernah putus, pengorbanan yang selalu diusahakan tanpa pamrih, dan kasih sayang yang tak terhingga sehingga penulis mampu menyelesaikan skripsi ini. Skripsi ini adalah bukti dari banyaknya pengorbanan papa dan mama. Harapan yang selalu kalian genggam adalah semangat bagi penulis dalam menyelesaikan semua hal yang sudah penulis mulai.
7. Saudara kandung penulis Gilang Joy Habib Ramadhan dan Faizah Fairuz Al-Arifi, terima kasih atas dukungan dan semangat yang diberikan secara tidak langsung untuk penulis dalam menyelesaikan skripsi ini.
8. Sahabat penulis Annisa Oktaviani, terima kasih sudah menjadi seseorang yang bisa diajak bertukar ide, mendengarkan keluh kesah penulis dengan baik, dan memberikan semangat untuk penulis agar tidak menyerah terhadap semua hal yang sudah dimulai.
9. Keluarga kedua di perantauan Cici, Riska, Nazla, Khumaidi, terima kasih sudah selalu ada dan mau direpotkan dalam hal apapun, terima kasih atas dukungan dan motivasi kalian.
10. Mas Nur Huda dan Rahmat Albar yang juga berkontribusi besar dan membantu penulis untuk tetap memahami, mengirimkan support, doa, dan semangat untuk tidak menyerah selama proses penulisan skripsi ini.

Semoga segala kontribusi, bantuan, dan dukungan yang diberikan kepada peneliti menjadi amal ibadah yang diterima oleh Allah SWT serta mendapatkan ridha-Nya. Peneliti telah berupaya semaksimal mungkin dalam menyusun skripsi ini dengan penuh dedikasi dan ketelitian. Namun, peneliti menyadari bahwa karya ini masih jauh dari kata sempurna. Oleh karena itu, dengan kerendahan hati, peneliti sangat mengharapkan saran dan masukan yang membangun dari berbagai pihak guna penyempurnaan skripsi ini di masa mendatang. Semoga skripsi ini memberikan manfaat bagi para pembaca, khususnya dalam pengembangan ilmu pengetahuan serta peningkatan kualitas pembelajaran.

Padang, 23 April 2026

Mutiara Salsabila Warman
NIM.21031025

DAFTAR ISI

	Halaman
ABSTRAK	i
<i>ABSTRACT</i>	ii
KATA PENGANTAR.....	iii
DAFTAR ISI	vi
DAFTAR TABEL.....	viii
DAFTAR LAMPIRAN	ix
BAB I. PENDAHULUAN	1
A. Latar Belakang Masalah.....	1
B. Identifikasi Masalah	7
C. Batasan Masalah.....	7
D. Rumusan Masalah	8
E. Tujuan Penelitian.....	8
F. Manfaat Penelitian	8
BAB II. KERANGKA TEORI.....	10
A. Kajian Teori.....	10
B. Penelitian yang Relevan.....	35
C. Kerangka Berpikir.....	38
BAB III. METODE PENELITIAN.....	39

A. Jenis Penelitian.....	39
B. Definisi Istilah.....	39
C. Tempat dan Waktu Penelitian.....	41
D. Subjek dan Objek Penelitian	41
E. Data Penelitian	41
F. Prosedur Penelitian.....	42
G. Instrumen Penelitian.....	43
H. Teknik Analisis Data	43
BAB IV. HASIL DAN PEMBAHASAN.....	49
A. Hasil Penelitian	49
B. Pembahasan.....	56
BAB V. PENUTUP	63
A. Kesimpulan	63
B. Saran.....	63
DAFTAR PUSTAKA	65
LAMPIRAN.....	73

DAFTAR TABEL

Tabel	Halaman
1. Kategori Kelayakan Soal Berdasarkan Persentase Validasi	48
2. Kualitas Item Kelas X	50
3. Kualitas Item Kelas XI	52
4. Kualitas Item Kelas XII	53
5. Data Demografi	54

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Lembar Angket Observasi.....	73
2. Soal Biologi SMA Kelas X Buatan Gemini AI.....	81
3. Soal Biologi SMA Kelas XI Buatan Gemini AI	92
4. Soal Biologi SMA Kelas XI Buatan Gemini AI	102
5. Lembar Kisi Kisi Soal Biologi SMA Kelas X	111
6. Lembar Kisi Kisi Soal Biologi SMA Kelas XI	128
7. Lembar Kisi Kisi Soal Biologi SMA Kelas XII.....	146
8. Lembar Validasi Logis Soal Kelas X Oleh Ahli 1.....	193
9. Lembar Validasi Logis Soal Kelas XI Oleh Ahli 1	197
10. Lembar Validasi Logis Soal Kelas XII Oleh Ahli 1	201
11. Lembar Validasi Soal Kelas X Oleh Ahli 2.....	205
12. Lembar Validasi Logis Soal Kelas XI Oleh Ahli 2	207
13. Lembar Validasi Logis Soal Kelas XII Oleh Ahli 2.....	209
14. Wright Maps Kelas X.....	211
15. Wright Map Kelas XI.....	212
16. Wright Map Kelas XII.....	213
17. Tabel Misfit Order Kelas X.....	214
18. Tabel Misfit Order Kelas XI	215
19. Tabel Misfit Order Kelas XII	216
20. Tabel Measure Order Kelas X.....	217
21. Tabel Measure Order Kelas XI	218

22. Tabel Measure Order Kelas XII	219
23. Summary Statistic Kelas X, XI dan XII.....	220
24. Tabel Item Distraktor Kelas X	221
25. Tabel Item Distraktor Kelas XI	224
26. Tabel Item Distraktor Kelas XII.....	226
27. Tabel Item DIF Jenis Kelamin Kelas X	228
28. Grafik Bias Item DIF Jenis Kelamin.....	229
29. Tabel Item DIF Jenis Kelamin Kelas XI	230
30. Tabel Item DIF Jenis Kelamin Kelas XII.....	231
31. Tabel Item DIF Jumlah Saudara Kelas X.....	232
32. Tabel Item DIF Jumlah Saudara Kelas XI	233
33. Tabel Item DIF Jumlah Saudara Kelas XII	234
34. Tabel Item DIF Pekerjaan Orang Tua Kelas X	235
35. Tabel Item DIF Pekerjaan Orang Tua Kelas XI	236
36. Tabel Item DIF Pekerjaan Orang Tua Kelas XII.....	237
37. Surat Izin Observasi FMIPA	238
38. Surat Izin Observasi Kemendikbud	239
39. Surat izin penelitian FMIPA.....	240
40. Surat Izin Penelitian Kemendikbud	241
41. Dokumentasi	242

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Pembelajaran biologi di tingkat SMA berperan penting dalam membentuk pemahaman ilmiah peserta didik serta mengembangkan kemampuan berpikir kritis, logis, dan analitis. Selain menyampaikan konsep dan prinsip dasar, biologi juga melatih keterampilan siswa dalam mencari informasi dan menerapkan metode ilmiah untuk memecahkan masalah (Jayanti, 2021). Kurikulum Merdeka mengarahkan pembelajaran biologi agar lebih fleksibel, kontekstual, dan sesuai dengan kebutuhan peserta didik dalam menghadapi tantangan abad ke-21 (Rochim *et al.* , 2024). Guru diberi keleluasaan untuk menyusun pembelajaran berdasarkan potensi siswa dan kondisi sekolah, sehingga proses belajar menjadi lebih adaptif (Sandy & Nugrahaningsih, 2024). Oleh karena itu, evaluasi yang relevan diperlukan untuk mendukung pembelajaran yang efektif, karena berfungsi sebagai alat ukur keberhasilan dalam mencapai tujuan pembelajaran biologi secara optimal.

Evaluasi pembelajaran memegang peranan strategis dalam pendidikan, termasuk dalam pembelajaran biologi. Evaluasi berfungsi untuk mengukur tingkat pencapaian tujuan pembelajaran sekaligus menjadi dasar dalam merancang strategi peningkatan mutu proses belajar mengajar (Rahmi, 2016). Dalam implementasi Kurikulum Merdeka, evaluasi tidak hanya berorientasi pada hasil akhir, tetapi juga menilai proses pembelajaran secara menyeluruh dengan memperhatikan perkembangan aspek kognitif, afektif, dan psikomotorik (Artama *et al.* , 2023).

Evaluasi yang dilaksanakan secara sistematis, berkelanjutan, dan berbasis pendekatan autentik mampu memberikan gambaran utuh mengenai kemampuan peserta didik. Hal ini penting sebagai acuan bagi guru dalam menentukan langkah tindak lanjut pembelajaran (Wisman *et al.* , 2021), serta dalam mengevaluasi dan menyempurnakan proses pembelajaran yang telah berlangsung (Rosyadi, 2024). Dengan demikian, evaluasi memiliki peran krusial dalam memastikan efektivitas pembelajaran dan mendukung pencapaian tujuan pendidikan secara holistik.

Proses evaluasi pembelajaran dapat berjalan secara terarah dan efektif apabila didukung dengan penggunaan alat atau metode penilaian yang tepat. Dalam pendidikan, alat evaluasi berfungsi mengumpulkan informasi terkait pencapaian belajar peserta didik, yang dapat berupa instrumen tes maupun non-tes (Septiana, 2016). Tes adalah bentuk evaluasi melalui pemberian tugas atau serangkaian soal baik pertanyaan maupun instruksi untuk memperoleh nilai sebagai indikator kemampuan siswa (Rahman *et al.* , 2019). Sementara itu, metode non-tes memungkinkan penilaian dilakukan tanpa pengujian langsung, melainkan dengan teknik seperti observasi sistematis, wawancara, atau analisis dokumen relevan (Asrul *et al.* , 2022). Dengan demikian, pemilihan alat evaluasi yang sesuai sangat penting untuk menjamin keakuratan data yang diperoleh dalam menilai hasil belajar.

Instrumen evaluasi pada pembelajaran Biologi seharusnya dirancang secara sistematis serta selaras dengan tujuan pembelajaran agar mampu mengukur aspek kognitif, afektif, dan psikomotor peserta didik secara tepat. Bentuk tes evaluasi yang umum digunakan meliputi tes tertulis seperti pilihan ganda, benar-salah,

menjodohkan, isian singkat, dan uraian, serta tes lisan dan tes perbuatan atau praktik (Usman & Mustaqimah, 2023). Penyusunan instrumen evaluasi perlu mempertimbangkan prinsip triangulasi antara tujuan pembelajaran, kegiatan pembelajaran, dan evaluasi. Prinsip ini menekankan bahwa setiap instrumen evaluasi harus mencerminkan keterkaitan langsung dengan tujuan dan aktivitas pembelajaran yang telah dilaksanakan (Arikunto, 2011).

Penilaian yang disusun secara terarah dan sesuai dengan capaian pembelajaran sangat penting dilakukan oleh pendidik. Ketidaksesuaian antara bentuk soal dengan tujuan pembelajaran dapat menimbulkan ketidakakuratan dalam hasil evaluasi dan tidak mencerminkan capaian sebenarnya dari peserta didik (Lodang & Bara, 2012). Penilaian yang baik tidak hanya menunjukkan sejauh mana siswa menguasai materi, tetapi juga dapat mengungkap kesulitan belajar serta posisi siswa dalam kelompoknya (Widianingrum & Sudarmin, 2014). Dengan demikian, hasil evaluasi dapat dijadikan dasar dalam pengambilan keputusan dan perbaikan proses pembelajaran.

Soal biologi yang baik harus memenuhi standar kualitas yang meliputi aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Soal yang baik harus sesuai dengan aspek isi menggunakan bahasa yang komunikatif dan sesuai dengan kaidah Bahasa Indonesia, serta tidak menimbulkan ambiguitas bagi siswa (Sapitri *et al.* , 2022). Soal juga harus mengacu pada indikator yang telah ditetapkan, serta mencakup berbagai level kognitif, mulai dari mengingat hingga menganalisis (Jannah *et al.* , 2023). Selain itu, proporsi tingkat kesulitan soal harus seimbang agar dapat mengukur kemampuan siswa secara

menyeluruh; misalnya, soal mudah, sedang, dan sulit harus tersebar secara proporsional (Rajagukguk & Naibaho, 2023). Efektivitas pengecoh dalam soal pilihan ganda juga menjadi faktor penting untuk menghindari jawaban yang mudah ditebak dan meningkatkan kualitas soal secara keseluruhan (Oktanin, 2023). Dengan memenuhi standar-standar tersebut, soal biologi dapat memberikan gambaran yang akurat tentang penguasaan materi siswa dan mendukung proses pembelajaran yang bermakna.

Berdasarkan hasil wawancara yang dilakukan peneliti bersama Ibu Dean Roslaini S.Pd, M.M, selaku guru mata pelajaran Biologi di SMA Pembangunan Laboratorium UNP Kota Padang pada tanggal 10 April 2025, ditemukan beberapa kendala yang dihadapi guru dalam menyusun soal evaluasi pembelajaran. Salah satu tantangan utama dalam pelaksanaan evaluasi pembelajaran adalah terbatasnya ragam soal yang tersedia dalam bank soal SMA Pembangunan Laboratorium UNP, sehingga materi evaluasi yang digunakan cenderung bersifat repetitif dan kurang mampu menggambarkan cakupan kompetensi secara menyeluruh. Di samping itu, alokasi waktu yang dimiliki guru untuk merancang instrumen evaluasi sering kali tidak mencukupi, sehingga menghambat proses pengembangan soal yang inovatif dan sesuai dengan kebutuhan peserta didik. Upaya pemanfaatan teknologi sebenarnya telah mulai diterapkan dalam penyusunan soal, namun penerapannya masih belum berjalan secara maksimal karena guru mengalami kesulitan dalam mengakses dan mengimplementasikan teknologi secara berkelanjutan dalam kegiatan evaluasi.

Di era *society* 5.0, tantangan dalam dunia pendidikan semakin kompleks. Guru dituntut tidak hanya untuk menyampaikan materi, tetapi juga meningkatkan kualitas sumber daya manusia melalui inovasi dalam proses pembelajaran (Subandowo, 2022). Salah satu upaya yang dapat dilakukan adalah dengan menyediakan soal evaluasi yang berkualitas dan relevan bagi peserta didik (Rahmadhani, 2014). Pesatnya perkembangan teknologi informasi dan komunikasi telah mengubah cara manusia belajar dan berinteraksi. Oleh karena itu, integrasi teknologi berbasis kecerdasan buatan atau *Artificial Intelligence* (AI) dapat menjadi solusi strategis untuk mengatasi kendala dalam penyusunan soal evaluasi.

Kemajuan teknologi kecerdasan buatan telah membawa dampak signifikan dalam berbagai bidang, termasuk pendidikan, khususnya dalam pembuatan soal evaluasi. Salah satu web AI, yaitu Gemini AI, menawarkan solusi inovatif untuk membantu guru dalam menyusun perangkat ajar, materi ajar, serta membuat soal yang bervariasi dengan lebih efisien (Aprilia *et al.* , 2024). Gemini AI, sebagai salah satu AI generatif multimodal, memiliki kemampuan canggih dalam memahami dan menghasilkan teks, gambar, dan data dalam berbagai format, yang memungkinkan guru untuk menciptakan pembelajaran yang interaktif yang mendukung berbagai karakter siswa (FIP Unesa, 2022).

Pemanfaatan teknologi kecerdasan buatan, seperti Gemini AI, dalam dunia pendidikan menjadi salah satu terobosan yang potensial untuk meningkatkan efektivitas penyusunan materi ajar serta instrumen evaluasi pembelajaran yang inovatif dan adaptif terhadap kebutuhan zaman. Penelitian yang dilakukan oleh Suganda (2023) menunjukkan bahwa Gemini AI merupakan alternatif yang

menjanjikan untuk diintegrasikan ke dalam dunia pendidikan, khususnya dalam pengembangan materi ajar dan soal evaluasi. Kemampuannya dalam menghasilkan teks kreatif dalam berbagai format memberikan peluang bagi guru untuk menciptakan bahan ajar yang lebih menarik serta soal-soal yang bervariasi.

Beberapa penelitian terdahulu menunjukkan bahwa ketertarikan terhadap pemanfaatan kecerdasan buatan dalam pengembangan soal evaluasi didasari oleh keterbatasan penelitian yang secara langsung menguji kualitas butir soal yang dibuat oleh AI. Penelitian oleh Kasprianto *et al.*, (2025) secara eksplisit menyebutkan bahwa latar belakang dilakukannya penelitian mereka adalah belum adanya pengujian kualitas soal berbasis *High Order Thinking Skills (HOTS)* yang disusun menggunakan teknologi AI di lingkungan sekolah, sehingga mereka terdorong untuk melakukan kajian terhadap soal yang dikembangkan dengan bantuan AI. Sejalan dengan temuan Zuhri *et al.*, (2024) yang menyatakan bahwa meskipun kecerdasan buatan menawarkan berbagai kemudahan dan keunggulan dalam dunia pendidikan, penggunaannya tetap perlu disikapi secara cermat, khususnya dalam memastikan mutu butir soal yang digunakan dalam evaluasi.

Penelitian ini berangkat dari asumsi serupa: bahwa belum ada studi yang secara sistematis menguji kualitas butir soal yang dihasilkan oleh kecerdasan buatan. Ketiadaan validasi empiris ini membatasi pemahaman kita tentang keandalan dan efektivitas butir soal AI, sehingga mendorong dilakukannya penelitian ini untuk menjembatani kesenjangan tersebut. Senada dengan itu, penelitian lain yang dilakukan dalam konteks pendidikan biologi juga memperlihatkan bahwa penggunaan AI masih berada dalam tahap eksploratif, dan

belum banyak menyentuh aspek validasi kualitas instrumen evaluasi secara komprehensif. Dengan demikian, penelitian yang menitikberatkan pada pengujian kualitas butir soal Biologi yang dikembangkan melalui platform kecerdasan buatan, khususnya Web Gemini, masih memiliki relevansi yang tinggi untuk dilakukan, karena dapat mengisi kekosongan kajian yang hingga saat ini belum banyak ditelaah secara mendalam oleh penelitian sebelumnya. Maka dari itu, peneliti tertarik untuk melakukan penelitian mengenai uji kualitas soal Biologi SMA yang dikembangkan melalui Gemini AI.

B. Identifikasi Masalah

Berdasarkan latar belakang, beberapa masalah yang teridentifikasi adalah:

1. Guru Biologi SMA masih mengalami kesulitan dalam menyusun soal evaluasi yang bervariasi karena keterbatasan waktu yang dimiliki.
2. Pemanfaatan teknologi AI, terutama Gemini AI, sebagai salah satu solusi permasalahan dalam pembuatan soal evaluasi yang bervariasi masih belum digiatkan.
3. Belum terdapat penelitian yang menguji kualitas butir soal Biologi SMA yang dikembangkan menggunakan teknologi AI, seperti Gemini.

C. Batasan Masalah

Berdasarkan identifikasi masalah, penelitian ini memiliki batasan sebagai berikut:

1. Penelitian ini hanya menganalisis kualitas soal pilihan ganda Biologi tingkat SMA yang dikembangkan menggunakan Gemini AI, dengan fokus pada lima aspek kualitas soal berdasarkan analisis pemodelan Rasch; validitas, reliabilitas,

indeks separasi, tingkat kesukaran soal, item distraktor, dan data demografi (item DIF).

D. Rumusan Masalah

Berdasarkan batasan masalah, penelitian ini memiliki rumusan masalah sebagai berikut:

Bagaimana kualitas soal pilihan ganda Biologi tingkat SMA yang dikembangkan menggunakan Gemini AI yang dianalisis dengan Rasch Model?

E. Tujuan Penelitian

Berdasarkan rumusan masalah, penelitian ini memiliki tujuan untuk menganalisis kualitas soal pilihan ganda Biologi tingkat SMA yang dikembangkan menggunakan Gemini AI dengan analisis Rasch Model.

F. Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

- a. Manfaat Teoretis
 1. Memberikan kontribusi pada pengembangan ilmu evaluasi pembelajaran, khususnya analisis kualitas butir soal berbantuan teknologi kecerdasan buatan (AI).
 2. Menambah literatur mengenai pemanfaatan AI dalam bidang pendidikan, terutama dalam penyusunan instrumen evaluasi yang valid, reliabel, dan sesuai standar.
 3. Menjadi dasar referensi bagi penelitian lanjutan terkait integrasi AI dengan asesmen pendidikan di Indonesia.

b. Manfaat Praktis

1. Bagi Guru

Memberikan informasi tentang potensi dan keterbatasan penggunaan AI dalam mengembangkan soal evaluasi dengan level kognitif yang sesuai.

2. Bagi Pengembang AI

Memberikan masukan untuk meningkatkan kemampuan AI dalam merancang soal evaluasi dengan level kognitif yang akurat.

3. Bagi Siswa

Diharapkan dapat meningkatkan kualitas soal evaluasi yang digunakan, sehingga mampu mengukur pemahaman siswa secara lebih komprehensif.

4. Bagi Peneliti Lain

Memberikan referensi dan inspirasi untuk penelitian selanjutnya dalam bidang pengembangan soal evaluasi berbasis AI.

BAB V

PENUTUP

A. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan pada bab sebelumnya, maka didapatkan kesimpulan bahwa soal evaluasi mata pelajaran Biologi SMA yang dihasilkan oleh Gemini AI yang dianalisis menggunakan Rasch Model memiliki kualitas yang baik dengan catatan perbaikan pada indikator kualitas soal tertentu.

B. Saran

Berdasarkan hasil penelitian yang telah dilakukan, dapat disarankan hal-hal sebagai berikut:

1. Disarankan bagi guru untuk tidak hanya mengandalkan hasil soal yang dihasilkan oleh AI, tetapi tetap melakukan proses validasi dan revisi, terutama pada butir soal yang tidak fit, memiliki distraktor tidak berfungsi, atau terindikasi bias, agar instrumen evaluasi lebih akurat dan adil.
2. Penggunaan analisis Model Rasch perlu terus dioptimalkan dalam mengevaluasi kualitas soal, karena mampu memberikan informasi yang lebih komprehensif terkait karakteristik butir soal dan kemampuan peserta didik.
3. Penelitian selanjutnya disarankan untuk menggunakan jumlah sampel yang lebih besar dan beragam agar hasil analisis, khususnya pada aspek reliabilitas dan DIF, menjadi lebih stabil dan dapat digeneralisasikan.
4. Penelitian berikutnya juga perlu mengembangkan variasi bentuk instrumen evaluasi, tidak hanya soal pilihan ganda, tetapi juga soal uraian atau asesmen

berbasis kinerja untuk memperoleh gambaran kemampuan siswa yang lebih komprehensif.

5. Disarankan untuk mengkaji lebih lanjut pengaruh kualitas *prompt* (seperti teknik *RISE*) terhadap hasil soal yang dihasilkan AI, serta melakukan perbandingan dengan soal buatan guru untuk mengetahui efektivitas dan kelayakan penggunaan AI dalam evaluasi pembelajaran.

DAFTAR PUSTAKA

- Afrita, J. (2023). Peran artificial intelligence dalam meningkatkan efisiensi dan efektifitas sistem pendidikan. *COMSERVA: Jurnal Penelitian Dan Pengabdian Masyarakat*, 2(12), 3181-3187.
- Ali, M. K., & Hasanah, A. (2024). Efektivitas fitur chatgpt, gemini dan claude ai dalam membantu guru membuat bahan ajar. *PEDAGOGIC: Indonesian Journal of Science Education and Technology*, 4(1), 58-71.
- Alti, R. P., Zulyuri, Z., & Violita, V. (2022). Analisis Kualitas Butir Soal Ujian Tengah Semester Mata Pelajaran Biologi Kelas X di MAN 1 Solok Selatan. *Jurnal Metaedukasi: Jurnal Ilmiah Pendidikan*, 4(2), 70-75.
- Andrich, D. (2018). *Controlling response dependence in the measurement of change using the Rasch model. Statistical Methods in Medical Research*, 27(12), 3709-3725.
- Anwar, Z., Masrifatin, Y., & Rahmania, N. (2024). Analisis Butir Soal Penilaian Akhir Semester MI Menggunakan Rasch Model. *Mentari: Journal of Islamic Primary School*, 2(2), 151-161.
- Aprilia, L. R., Istifarah, A. L., & Prayitno, S. (2024). Efektivitas Penggunaan Gemini AI dalam Menyusun Perangkat Ajar bagi Guru. *Seminar Nasional Teknologi Informasi dan Komunikasi (SeNTIK STI&K)*, 257-261. <https://ejournal.jak-stik.ac.id/index.php/sentik/article/view/3728/857>
- Arifin, Z. (2009). *Evaluasi pembelajaran* (Vol. 8). Bandung: Remaja Rosdakarya.
- Arikunto, S. (2011). *Dasar-dasar Evaluasi Pendidikan* (Edisi revisi). Jakarta: Bumi Aksara.
- Artama, S., Djollong, A. F., Ismail, I., Lubis, L. H., et al. (2023). *Evaluasi Hasil Belajar*. Medan: PT. Mifandi Mandiri Digital.
- Aslamiya, P. Z. (2024). *Analisis butir soal fisika berbasis kearifan lokal tari tanggai buatan artificial intelligence pada materi dinamika gerak* (Skripsi Sarjana, Universitas Sriwijaya). Universitas Sriwijaya Repository. <https://repository.unsri.ac.id/146007/>
- Asrul, A., Saragih, A. H., & Mukhtar, M. (2022). *Evaluasi pembelajaran*. Perdana Publishing.
- Azizah, A., & Wahyuningsih, S. (2020). Penggunaan model RASCH untuk analisis instrumen tes pada mata kuliah matematika aktuaria. *Jupitek*, 3(1), 45-50.
- Barel, E., dan Tzischinsky, O. 2018. *Age and Sex Differences in Verbal and Visuospatial Abilities. Advances in Cognitive Psychology*, 14 (2), 51-61.

- Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.
- Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how?. *CBE—Life Sciences Education*, 15(4), rm4.
- Briggs, D. C. (2019). Interpreting and visualizing the unit of measurement in the Rasch Model. *Measurement*, 146, 961-971.
- Budi, A. S., Wulandari, A., & Syahroni, M. (2023). Kualitas Soal Penilaian Tengah Semester Mata Pelajaran Bahasa Indonesia Jenjang Sekolah Menengah Pertama. *Indonesian Journal of Education and Learning*, 7(1).
- Burhan, B., Abdillah, S. A. S., Rizal, A., Muhammad, A. F., January, R., & Rampeng, R. (2025). Inovasi Pembelajaran Di Era Digital: Memperkuat Kompetensi Guru Melalui Gemini Artificial Intelligence (AI). *DEVOSI*, 6(1), 27-40.
- Cahyanto, H. N. (2025). *Metodologi Penelitian Dengan Ai (Artificial Intelligence)*. PUSTAKA SAGA.
- Darwis, N., & Efendi, R. (2024). Bentuk-Bentuk Tes Dan Karakteristiknya. *Jurnal Ilmiah Kajian Multidisipliner*, 8(12).
- Daryanto. (2008). *Evaluasi Pendidikan*. Jakarta: PT Rineka Cipta.
- Diantama, S. (2024). Pemanfaatan Artificial Intelegent (AI) Dalam Dunia Pendidikan. *Dewantech Jurnal Teknologi Pendidikan*, 2(1), 11-17.
- Farid, I., Reksoprodjo, A. H., & Suhirwan, S. (2023). Pemanfaatan Artificial Intelligence Dalam Pertahanan Siber. *Nusantara: Jurnal Ilmu Pengetahuan Sosial*, 10(2), 779-788.
- Fauziana, A., & Wulansari, A. D. (2021). Analisis Kualitas Butir Soal Ulangan Harian di Sekolah Dasar dengan Model Rasch. *Jurnal Ibriez: Jurnal Kependidikan Dasar Islam Berbasis Sains*, 6(1), 10-19.
- FIP Unesa. (2022). Implementasi Gemini AI dalam Pendidikan: Meningkatkan Pembelajaran dan Manajemen Akademik. Fakultas Ilmu Pendidikan, Universitas Negeri Surabaya. <https://fip.unesa.ac.id/implementasi-gemini-ai-dalam-pendidikan-meningkatkan-pembelajaran-dan-manajemen-akademik/>
- Fitri, L. A. (2019). *Analisis Soal Ujian Akhir Semester Buatan Guru Biologi Kelas XI IPA Semester Genap di SMAN 1 Kuta Cot Glie Menggunakan Aplikasi Anates Versi 4.09* (Doctoral dissertation, UIN Ar-Raniry).
- Haladyna, T. M., & Rodriguez, M. C. (2021). Using full-information item analysis to improve item quality. *Educational Assessment*, 26(3), 198-211.

- Han, C. (2019). William J. Boone, John R. Staver and Melissa S. Yale. Rasch Analysis in the Human Sciences. *Journal of Research Design and Statistics in Linguistics and Communication Science*, 5(1-2), 208-211.
- Hasanein, A., Sobaih, A., & Elshaer, I. (2024). Examining Google Gemini's acceptance and usage in higher education. *Journal of Applied Learning and Teaching*, 7(2).
- Hendriani, M., Marfilinda, R., & Apfani, S. (2022). Pelatihan pembuatan soal evaluasi pembelajaran berbasis digital di era disrupsi teknologi. *Jurnal Abdi Insani*, 9(1), 247-255.
- Idrus, L. (2019). Evaluasi dalam proses pembelajaran. *Adaara: Jurnal Manajemen Pendidikan Islam*, 9(2), 920-935.
- Imran, M., & Almusharraf, N. (2024). *Google Gemini as a next generation AI educational tool: a review of emerging educational technology*. *Smart Learning Environments*, 11(1), 22.
- Iskandar, A., & Rizal, M. (2018). Analisis kualitas soal di perguruan tinggi berbasis aplikasi TAP. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 22(1), 12-23.
- Jannah, I. K., Mahanal, S., & Mashfufah, A. (2023). Analisis Tingkat Kognitif Soal Asesmen Sumatif Akhir Semester I (ASAS I) IPA Berbasis Jenis Soal AKM berdasarkan Taksonomi Bloom di Kelas V SD Swasta Kota Malang. *JIIP- Jurnal Ilmiah Ilmu Pendidikan*, 6(2), 806-810.
- Jayanti, U. N. A. D. (2021). *Perencanaan Pembelajaran Biologi: Tinjauan Teori, Praktik, dan Paradigma Wahdatul 'Ulum*. Medan: Perdana Publishing.
- Jayawardana, H. B. A., & Gita, R. S. D. (2020). Inovasi Pembelajaran Biologi Di Era Revolusi Industri 4.0. In *Prosiding Seminar Nasional Biologi* (Vol. 6, No. 1, pp. 58-66).
- Jumini, S., Madnasri, S., Cahyono, E., & Parmin, P. (2023, June). Analisis kualitas butir soal pengukuran literasi sains melalui teori tes klasik dan rasch model. In *Prosiding Seminar Nasional Pascasarjana* (Vol. 6, No. 1, pp. 758-765).
- Kasprianto, R., Munawar, W., & Sriyono, S. Analisis Butir Soal High Order Thinking Skills (HOTS) Berbantuan Artificial Intelligence (AI) Untuk Pembelajaran Perawatan Dan Perbaikan Sasis Sepeda Motor Di SMK. *ATIKANOTO: Journal of Automotive Engineering Education*, 2(1).
- Kristianti, T. (2023). Implementasi Artificial Intelligence (Ai) Dalam Dunia Pendidikan Di Era Society 5.0. *Prosiding Temu Ilmiah Nasional Guru*, 15(1), 145-155.
- Kurniawan, A., Febriant, A. N., Hardianti, T., Ichsan, I., Desy, D., Risan, R., ... & Fuad, H. (2022). Evaluasi pembelajaran.
- Limna, P., & Kraiwanit, T. (2024). *Assessing the determinants of generative AI integration: A study on Google's Gemini adoption among Thai University*

- students. *Journal of University Teaching and Learning Practice*, 21(10), 1-20.
- Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean. *Rasch measurement transactions*, 16(2), 878.
- Lodang, H., & Bara, N. A. S. (2012). Analisis Kesesuaian Antara Instrumen Evaluasi Formatif dengan Tujuan Kognitif Pembelajaran Biologi di SMP Watansoppeng. *Jurnal Bionature*, 13(2), 120–126.
- Maulani, G., Septiani, S., Susilowaty, N., Rusmayani, N. G. A. L., Sukamdi, S., Evenddy, S. S., Lasri, L., Nababan, H. S., Setiadi, K., Rahayu, I., Simanungkalit, L. N., Edi, S., Sari, N., Atin, S., Thoif, M., Raju, M. J., Saptadi, N. T. S., Rosalinda, R., Syafriansyah, M., & Nurlily, L. (2020). *Evaluasi pembelajaran*. Sada Kurnia Pustaka. <https://books.google.co.id/books?id=ebQVEQAAQBAJ>
- Muchamad, M. K., Khamdan, T. A. G., & Wahyuni, P. (Eds.). (2021). *Kampus Merdeka Seri 1: Menilik Kesiapan Teknologi Dalam Sistem Kampus*. Banda Aceh: Syiah Kuala University Press.
- Muchtar, H. (2010). Penerapan penilaian autentik dalam upaya peningkatan mutu pendidikan. *Jurnal Pendidikan Penabur*, 14(9), 68-76.
- Mulyani, S., et al. (2023). Analisis kualitas soal evaluasi pada buku ajar. *Jurnal Inspirasi Pendidikan ALFIHRIS*, 3(2), 45–53.
- Nazharita, E., Hidayat, R., & Rahmatudin, J. (2025). Analisis Butir Soal Tes Literasi Numerasi Domain Geometri Siswa SMP Menggunakan Rasch Model Berbantuan Ministep. *Pedagogy: Jurnal Pendidikan Matematika*, 10(3), 989-1000.
- Nofiana, M. (2015). Kualitas penulisan butir soal ujian nasional biologi tahun 2014/2015 ditinjau dari aspek teoritik. *Saintifika*, 17(1).
- Novriyanti, E., & Arthur, R. (2024). Analisis kualitas butir soal ujian tengah semester Biologi umum menggunakan Model Rasch. *JagoMIPA: Jurnal Pendidikan Matematika dan IPA*, 4(4), 718-733.
- Nurdyansyah, N., & Fahyuni, E. F. (2016). Inovasi model pembelajaran sesuai kurikulum 2013.
- Oktanin, A. (2023). Kualitas Butir Soal Biologi Kelas X IPA MAN 1 Merangin Tahun Ajaran 2022/2023. *EDU-BIO Jurnal Pendidikan Biologi*, 6(1), 17-40.
- Pebrian, Y., & Farhat, M. F. (2023). Pemanfaatan Artificial Intelligence Dalam Dunia Pendidikan. *Abdi Jurnal Publikasi*, 2(2), 84-87.
- Perera, P., & Lankathilaka, M. (2023). Preparing to revolutionize education with the multi-model GenAI tool Google Gemini? A journey towards effective

- policy making. *Journal of Advances in Education and Philosophy*, 7(8), 246–253
- Pujiono, I. P., Sopiah, S., Sofyan, N. H., & Arifin, J. (2024). Workshop Google Gemini Untuk Meningkatkan Pengetahuan Siswa-Siswi Di SMP Negeri 1 Kadangserang. *Jurnal Pengabdian Masyarakat Polmanbabel*, 4(02), 129–135.
- Qomari, R. (2008). Pengembangan instrumen evaluasi domain afektif. *INSANIA: Jurnal Pemikiran Alternatif Kependidikan*, 13(1), 87–109.
- Rahmadhani, E. K. (2014). Analisis Kualitas Butir Soal Pada Bank Soal Biologi Kelas X SMA. *Journal BioEdu*, 3(1), 422–438.
- Rahman, A. A., Pd Cut, M., Nasryah, E., & Pd, M. (2019). *EVALUASI PEMBELAJARAN. Uwais Inspirasi Indonesia*. <https://repository.bbg.ac.id/handle/837>
- Rahmi. (2016). Evaluasi Proses Pembelajaran Biologi Kelas I di SMPN Kabupaten Bireuen. *Jurnal Pendidikan Almuslim*, IV(1), 34–35.
- Rajagukguk, M. J. T., & Naibaho, D. (2023). Mampu Memilih Soal Berdasarkan Tingkat Kesukaran. *Jurnal Pendidikan Sosial Dan Humaniora*, 2(4), 12736–12747.
- Ramadhan, A. F., & Hidayatullah, R. S. (2023). Analisis kualitas butir soal ujian satuan pendidikan (USP) materi C2 teknik pemesinan kelas XII di SMK PGRI 1 Lamongan melalui model Rasch.
- Riduwan. (2015). *Skala Pengukuran Variabel-variabel Penelitian*. Bandung: Alfabeta.
- Rochim, A. A. (2024). Kecerdasan Buatan: Resiko, Tantangan Dan Penggunaan Bijak Pada Dunia Pendidikan. *Antroposen: Journal of Social Studies and Humaniora*, 3(1), 13–25.
- Rochim, A. A., Baharung, S., & Isnaini, I. (2024). Perencanaan Pembelajaran Biologi Berbasis Project Based Learning pada Kurikulum Merdeka di SMAN 1 Bungku Tengah. *Jurnal Pendidikan Tambusai*, 8(2), 18154–18161.
- Rosyadi, A. (2024). Evaluasi Tingkat Pemahaman Konsep Siswa pada Materi Komponen Ekosistem dan Interaksi Antar Komponen Kelas X SMA Negeri 2 Bayan Kabupaten Lombok Utara: (*Evaluation of Students' Concept Understanding Level on Ecosystem Components and Interactions between Components of Class X SMA Negeri 2 Bayan, North Lombok Regency*). *Otus Education: Jurnal Biologi dan Pendidikan Biologi*, 2(2), 89–102.
- Sandy, D. A., & Nugrahaningsih, W. H. (2024). Implementasi Kurikulum Merdeka pada Mata Pelajaran Biologi di Sekolah Daerah Rural. *Prosiding Semnas Biologi XII Tahun 2024 FMIPA Universitas Negeri Semarang*, 69–71.

- Sapitri, A., Kurniati, T., & Yuliawati, A. (2022). Analisis Kualitas Soal UAS Biologi SMA Kelas X dan XI MIA. *Bioeduca: Journal of Biology Education*, 4(1), 45–56.
- Septiana, N. (2016). Analisis butir soal ulangan akhir semester (UAS) biologi tahun pelajaran 2015/2016 kelas X dan XI pada MAN Sampit. *Edu Sains: Jurnal Pendidikan Sains dan Matematika*, 4(2).
- Sholehurrohman, R., Ilman, I. S., Pambudi, A., Afdhaludin, M., Citra, E. E., & Badiwibowo, S. (2025). Peningkatan Kompetensi Guru Teknologi Informasi di Lampung Selatan melalui Pelatihan *Prompt* Kecerdasan Buatan. *JPM (Jurnal Pemberdayaan Masyarakat)*, 10(2), 247-256.
- Suaedi, M. (2016). Analisis terhadap kualitas butir soal buatan guru IPA kelas VIII MTs Negeri Slawi tahun pelajaran 2015/2016. *Jurnal Kamil*, 1 (3), 1-15.
- Subandowo, M. (2022). Teknologi Pendidikan di Era Society 5.0. *Jurnal Sagacious*, 9(1).
- Subari, A., Lufri, L., & Syamsurizal, S. (2021). Analisis butir soal ujian akhir semester ganjil biologi kelas XI MAN 2 Kota Jambi. *Bioscientist: Jurnal Ilmiah Biologi*, 9(1), 45-53.
- Sudijono, A. (2011). *Pengantar Evaluasi Pendidikan*. Jakarta: PT Rajagrafindo Persada.
- Suganda, A. (2023). Memilih AI Yang Tepat Untuk Guru: Perbandingan Fitur Gemini, ChatGPT, dan Claude AI. *Jurnal Inovasi Teknologi Dan Edukasi Teknik*, 3(11), 1–10. <https://doi.org/10.17977/um084.v3.i11.2023.2>
- Sugiyono. 2013. *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Penerbit Alfabeta
- Suharsimi Arikunto. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Sulistyawati, W., Wahyudi, W., & Trinuryono, S. (2022). Analisis motivasi belajar siswa dengan menggunakan model pembelajaran blended learning saat pandemi covid-19 (deskriptif kuantitatif di SMAN 1 babadan Ponorogo). *Kadikma*, 13(1), 68-73.
- Sumardi, M. (2020). *Teknik pengukuran dan penilaian hasil belajar*. Yogyakarta: Deepublish Publisher.
- Sumarna Surapranata. (2005). *Analisis, Validitas, Reliabilitas dan Intrepretasi Hasil Tes*. Bandung: PT Remaja Rosdakarya.
- Sumintono, B. & Widhiarso W. 2015. *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*. Cimahi: Trim Komunikata.

- SURBAKTI, I. W. (2025). Analisis Kualitas Butir Soal Pada Uji Coba Evaluasi Pembelajaran Matematika. *Cendekia: Jurnal Ilmu Pengetahuan*, 5(1), 116-125.
- Taruklimbong, E. S. W., & Sihotang, H. (2023). Peluang dan tantangan penggunaan AI (Artificial Intelligence) dalam pembelajaran kimia. *Jurnal Pendidikan Tambusai*, 7(3), 26745-26757.
- Taylor, C. L. (2025). Best Practices for Leveraging PISA CT Data to Understand Gender Differences in Creative Thinking. *The Journal of Creative Behavior*, 59(2), e70034.
- Usman, N. F., & Mustaqimah, N. (2023). Analisis Penggunaan Alat Evaluasi Pembelajaran Biologi untuk Menunjang Kegiatan Penilaian di SMA Negeri 1 Mananggu. *Jurnal Normalita*, 11(3), 524–530.
- Utami, R. N. (2020). Analisis Soal Tes Sumatif SMA Biologi Kelas XII Ditinjau dari Taksonomi Bloom Revisi (*Bachelor's thesis*, Jakarta: FITK UIN Syarif Hidayatullah Jakarta).
- Werdiningsih, G. (2015). Analisis Kualitas Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran Ekonomi Kelas XII IPS SMAN 2 Banguntapan Tahun Ajaran 2014/2015. *Skripsi. Universitas Negeri Yogyakarta*.
- Widianingrum, R., & Sudarmin. (2014). Evaluasi Hasil Belajar dalam Pembelajaran IPA. *Jurnal Pendidikan IPA Indonesia*, 3(2), 135–140.
- Widoyoko, E. P. (2009). Evaluasi program pembelajaran. *Yogyakarta: Pustaka Pelajar*, 238.
- Widoyoko, E. P. (2012). Teknik Penyusunan Instrumen Penelitian. *Yogyakarta: Pustaka Pelajar*
- Wisman, Y., Effrata, E., & Tutesa, T. (2021). Penerapan konsep instrumen evaluasi hasil belajar. *Jurnal Ilmiah Kanderang Tingang*, 12(1), 1-9.
- Yang, C. C., Totzek, J. F., Lepage, M., Lavigne, K. M., 2023. *Sex Differences In Cognition and Structural Covariance-based Morphometric Connectivity: Evidence From 28.000+ UK Biobank Participants. Cerebral Cortex*, 33 (10341-10354).
- Yusuf, F. W. (2024). Analisis Butir Soal Asesmen Sumatif Biologi Materi Perubahan Lingkungan Dengan Menggunakan Anates Pada Kelas X SMA. *LEARNING: Jurnal Inovasi Penelitian Pendidikan dan Pembelajaran*, 4(1), 126-135.
- Zahara, S. L., Azkia, Z. U., & Chusni, M. M. (2023). Implementasi Teknologi Artificial Intelligence (AI) dalam Bidang Pendidikan. *Jurnal Penelitian Sains Dan Pendidikan (JPSP)*, 3(1), 15-20.

- Zuhri, N. Z., Sopian, A., Sauri, S., & Nurbayan, Y. (2024). Analisis Validitas dan Reliabilitas Soal Bahasa Arab Melalui Website OpExams Pembuat Soal Berbasis AI. *Jurnal Pendidikan Modern*, 9(2), 87-91.
- Zuhri, N. Z., Syihabuddin, S., & Tatang, T. (2024). Analisis Validitas, Reliabilitas, dan Tingkat Kesukaran Soal Bahasa Arab Tingkat SMP Berbasis Artificial Intelligence (AI) melalui Platform QuestionWell. *Jurnal Pendidikan Dan Pembelajaran Indonesia (JPPI)*, 4(2), 693-704.