
Identifikasi Hasil Pengembangan Soal *Tryout* Mahasiswa Pendidikan Profesi Guru

Sumari, Siti Faizah*

Pendidikan Profesi Guru (PPG), Sekolah Pascasarjana, Universitas Negeri Malang

*Email: faizah.siti.pasca@um.ac.id

ABSTRACT

This study aims to analyze the quality of the tryout instrument for students of the Industrial Chemical Engineering Professional Education (PPG) Study Program through the validity and reliability test of the questions. The method used is descriptive quantitative with 29 students as research subjects. The instrument consists of 35 multiple-choice questions analyzed using Corrected Item-Total Correlation for validity and Cronbach's Alpha for reliability. The results showed that of the 35 questions, only about 20-25% met the validity requirements, while the rest were declared invalid. The reliability value of the first instrument was 0.565 which is included in the low-sufficient category. After item selection and leaving 11 questions, the reliability increased significantly to 0.804 which is included in the good category. This finding indicates that the quality of the instrument is determined more by the quality of the questions, not the number of questions. The second instrument is more valid and reliable so it can be used for learning evaluation, while the first instrument requires revision. In this case, the second instrument is the result of a selection of the first instrument, not two different instruments. Therefore, it is necessary to revise the tryout questions so that the questions developed are not too easy for PPG students in the Industrial Chemical Engineering Study Program.

Keywords: Development, Identification, Results, Teacher Professional Education (TPE) Students

ABSTRAK

Penelitian ini bertujuan untuk menganalisis kualitas instrumen tryout mahasiswa Program Studi Pendidikan Profesi Guru (PPG) Teknik Kimia Industri melalui uji validitas dan reliabilitas butir soal. Metode yang digunakan adalah kuantitatif deskriptif dengan subjek penelitian sebanyak 29 mahasiswa. Instrumen berupa 35 butir soal pilihan ganda yang dianalisis menggunakan Corrected Item-Total Correlation untuk validitas dan Cronbach's Alpha untuk reliabilitas. Hasil penelitian menunjukkan bahwa dari 35 butir soal, hanya sekitar 20-25% yang memenuhi syarat validitas, sedangkan sisanya dinyatakan tidak valid. Nilai reliabilitas instrumen pertama sebesar 0,565 yang termasuk kategori rendah-cukup. Setelah dilakukan seleksi item dan menyisakan 11 soal, reliabilitas meningkat signifikan menjadi 0,804 yang masuk dalam kategori baik. Temuan ini menunjukkan bahwa kualitas instrumen lebih ditentukan oleh mutu soal, bukan jumlah soal. Instrumen kedua lebih valid dan reliabel sehingga dapat digunakan untuk evaluasi pembelajaran, sementara instrumen pertama memerlukan revisi. Dalam hal ini instrumen kedua merupakan hasil seleksi dari instrumen pertama, bukan dua instrumen yang berbeda. Oleh karena itu perlu dilakukan revisi terhadap soal tryout agar soal yang dikembangkan tidak terlalu mudah bagi mahasiswa PPG Prodi Teknik Kimia Industri.

Kata Kunci: Hasil, Identifikasi, Pengembangan, Mahasiswa Pendidikan Profesi Guru (PPG)

PENDAHULUAN

Pendidikan tinggi, khususnya di bidang teknik kimia industri, memiliki peran strategis dalam menyiapkan sumber daya manusia yang kompeten, inovatif, dan adaptif terhadap perkembangan teknologi serta kebutuhan industri (Fullan, 2016). Pergeseran menuju era Revolusi Industri 4.0 menuntut perguruan tinggi untuk tidak hanya menekankan aspek kognitif, tetapi juga keterampilan berpikir kritis, pemecahan masalah, dan kemampuan mengintegrasikan ilmu pengetahuan dengan teknologi (Trilling & Fadel, 2009; OECD, 2019). Pendidikan teknik kimia dituntut untuk beradaptasi dengan kebutuhan industri global

melalui metode pembelajaran inovatif, namun penerapannya masih menghadapi tantangan seperti manajemen waktu, distribusi tugas tidak merata, dan variabilitas motivasi mahasiswa (Hidayatullah & Fahmi, 2025). Peningkatan partisipasi, motivasi, minat, dan hasil belajar mahasiswa pada semua jenjang pendidikan menjadi hal utama dalam proses pembelajaran kimia (Gomez-Coma et al., 2023)

Sedangkan dalam konteks pembelajaran di Pendidikan Profesi Guru (PPG) Program Studi (Prodi) Teknik Kimia Industri, pemberian *tryout* atau tes uji coba menjadi salah satu instrumen penting untuk mengevaluasi capaian pembelajaran mahasiswa sebelum menghadapi ujian sesungguhnya. *Tryout* tidak hanya berfungsi sebagai sarana latihan, tetapi juga sebagai media diagnostik untuk mengidentifikasi kelemahan dan kekuatan mahasiswa dalam memahami konsep dasar maupun penerapannya (Black & Wiliam, 1998; Rust, 2002). Hasil analisis *tryout* dapat memberikan gambaran mengenai tingkat penguasaan materi, kesiapan menghadapi ujian, serta efektivitas strategi pembelajaran yang digunakan dosen (Brown & Knight, 1994).

Selain itu, analisis hasil *tryout* juga relevan dalam mengukur capaian pembelajaran berbasis kompetensi yang menekankan pada penguasaan *higher order thinking skills* (HOTS). Menurut Anderson dan Krathwohl (2001), pengembangan keterampilan analitis, evaluatif, dan kreatif perlu dilatihkan melalui evaluasi yang berkesinambungan. Dengan demikian, data hasil *tryout* dapat dijadikan dasar untuk merancang intervensi pembelajaran yang lebih tepat sasaran, seperti penguatan konsep, pengayaan materi, maupun remediasi (Brookhart, 2010).

Hasil penelitian terdahulu, pada saat pembelajaran kimia di jenjang SMA Kendari menunjukkan bahwa pemahaman guru terbatas, kesulitan identifikasi kebutuhan siswa, waktu terbatas, dan kurangnya fleksibilitas kurikulum serta fasilitas (Kiki et al., 2025). Lebih lanjut, tingkat motivasi siswa pedesaan dalam pembelajaran kimia lebih rendah dibandingkan siswa perkotaan, terutama karena tantangan infrastruktur Teknologi Informasi dan Komunikasi (TIK), seperti akses internet yang terbatas dan komputer yang tidak memadai (Nsabayezu et al., 2025). Di sisi lain, integrasi TIK dalam bentuk AI generatif (GenAI) ke dalam pendidikan kimia menawarkan peluang untuk meningkatkan keterlibatan dan pengembangan keterampilan siswa dan guru, tetapi penggunaannya menimbulkan pertanyaan tentang memastikan akurasi pengetahuan dan mendorong pemikiran kritis (Yuriev et al., 2025).

Dari hasil penelitian terdahulu dapat diketahui bahwa kemampuan guru dalam pembelajaran kimia perlu ditingkatkan, tidak hanya kimia secara umum saja tetapi juga Teknik kimia industri. Guru perlu memiliki kemampuan pedagogi dalam setiap pembelajaran. Dengan adanya PPG ini diharapkan dapat mencetak guru yang profesional dan mampu menguasai keterampilan pedagogi dan berpikir kritis. Kemampuan tersebut dapat diidentifikasi melalui *tryout* dan tes.

Lebih jauh, penggunaan hasil *tryout* dalam konteks pendidikan tinggi juga mendukung pencapaian *Sustainable Development Goals* (SDGs), khususnya tujuan keempat mengenai pendidikan berkualitas, serta sejalan dengan kebijakan Merdeka Belajar yang menekankan pada pembelajaran inovatif, adaptif, dan relevan dengan kebutuhan zaman (Kemendikbud, 2020). Oleh karena itu, analisis hasil *tryout* mahasiswa PPG Teknik Kimia Industri menjadi penting untuk mengetahui tingkat pencapaian kompetensi, mengidentifikasi tantangan pembelajaran, serta merumuskan strategi perbaikan yang dapat meningkatkan mutu pendidikan di perguruan tinggi. Oleh karena itu, tujuan penelitian ini adalah untuk mengidentifikasi hasil pengembangan soal *tryout* mahasiswa PPG Prodi Teknik Kimia Industri untuk menyiapkan strategi pembelajaran mahasiswa.

METODE

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan tujuan untuk menganalisis hasil *tryout* mahasiswa PPG Teknik Kimia Industri. Analisis dilakukan melalui uji validitas dan reliabilitas butir soal, sehingga dapat diketahui sejauh mana instrumen layak digunakan sebagai alat evaluasi pembelajaran (Creswell, 2012). Subjek penelitian adalah mahasiswa Program Profesi Guru (PPG) Universitas Negeri Malang Prodi Teknik Kimia Industri. Jumlah responden sebanyak 29 mahasiswa yang dipilih secara total sampling karena seluruh populasi peserta *tryout* dijadikan sampel penelitian. Jumlah tersebut sesuai dengan banyaknya mahasiswa PPG Prodi Teknik Kimia Industri yang sedang menempuh studi di Universitas Negeri Malang sehingga seluruh mahasiswa dijadikan sampel penelitian.

Instrumen yang digunakan berupa tes objektif sebanyak 35 butir soal pilihan ganda yang disusun untuk mengukur penguasaan kompetensi mahasiswa. Selanjutnya, instrumen ini diuji validitas dan reliabilitas. Data pada penelitian dikumpulkan dari hasil pengerjaan tes *tryout* mahasiswa. Setiap jawaban mahasiswa dikodekan dalam bentuk skor (1=benar, 0=salah). Hasil pengerjaan tes kemudian direkapitulasi dalam lembar kerja (*spreadsheet*) untuk dianalisis lebih lanjut.

Adapun prosedur dan analisis data penelitian terdiri dari beberapa langkah. Pertama, perencanaan penelitian yang meliputi: (a) Mengidentifikasi permasalahan terkait kualitas instrumen evaluasi pada mahasiswa PPG Prodi Teknik Kimia Industri; (b) menganalisis validitas dan reliabilitas butir soal *tryout* mahasiswa; (c) Menyusun rancangan penelitian dengan pendekatan kuantitatif deskriptif; Langkah kedua, penyusunan instrumen *tryout* berdasarkan kompetensi dasar (KD) dan indikator pencapaian kompetensi yang ada di RPS; (d) Mengembangkan instrumen evaluasi berupa 35 butir soal pilihan ganda yang dapat mengukur pemahaman mahasiswa; (c) Melakukan validasi; Langkah ketiga adalah pelaksanaan *tryout* kepada subjek penelitian yaitu seluruh mahasiswa PPG Prodi Teknik Kimia Industri yang terdiri dari 29 mahasiswa; Langkah keempat adalah analisis validitas butir soal yang meliputi: (a) Memasukkan data hasil tes ke dalam program SPSS versi 25; (b) Menghitung Corrected Item-Total Correlation untuk setiap butir soal; (c) Menentukan status validitas (*Valid* jika nilai r hitung $\geq 0,30$, *Tidak valid* jika nilai r hitung $< 0,30$ (Azwar, 2012), Mengidentifikasi butir soal yang valid untuk dipertahankan dan butir soal yang tidak valid untuk direvisi atau dieliminasi. Langkah kelima adalah analisis reliabilitas instrumen yang meliputi: (a) Menghitung nilai Cronbach's Alpha menggunakan SPSS; (b) Menafsirkan koefisien reliabilitas sesuai kriteria yakni: $\alpha \geq 0,90$:Sangat baik; $\alpha \geq 0,80$:Baik; $\alpha \geq 0,70$:Cukup; $\alpha \geq 0,60$:dipertanyakan; $\alpha < 0,50$:Rendah (George & Mallery, 2003). Langkah keenam adalah penyusunan hasil analisis yang meliputi: (a) Menyajikan hasil uji validitas dalam bentuk tabel distribusi butir valid dan tidak valid; (b) Menyajikan hasil uji reliabilitas dalam bentuk nilai Cronbach's Alpha dan kategori reliabilitas; (c) Menyusun deskripsi hasil analisis yang memuat kekuatan, kelemahan, serta rekomendasi perbaikan instrumen. Langkah terakhir, penarikan kesimpulan berdasarkan hasil penelitian dan pembahasan.

HASIL DAN PEMBAHASAN

Berdasarkan hasil analisis validitas soal diketahui bahwa dari 35 butir soal, sebagian besar tidak valid karena memiliki nilai korelasi item-total di bawah 0,30. Hanya sebagian kecil soal yang valid, yakni nomor 2, 3, 5, 9, 17, 20, 21, 23, dan 29. Hal ini menunjukkan bahwa hanya sekitar 25% soal yang benar-benar mampu mengukur konstruk pengetahuan yang diinginkan, sementara sisanya perlu direvisi atau dieliminasi. Sedangkan dari hasil analisis daya pembeda soal diketahui bahwa daya beda soal bervariasi, mayoritas berada pada kategori jelek (0,00–0,20). Hanya sedikit soal yang mencapai kategori cukup hingga sangat baik, seperti soal nomor 9 (0,500), nomor 20 (0,333), nomor 21 (0,500), dan nomor 23 (0,333). Kondisi ini mengindikasikan bahwa banyak soal yang tidak dapat digunakan untuk membedakan mahasiswa yang berkemampuan tinggi dan rendah.

Pada aspek tingkat kesukaran soal menunjukkan bahwa sebagian besar soal berada pada kategori mudah (indeks kesukaran $> 0,70$). Beberapa soal berada pada kategori sedang, misalnya soal nomor 5 (0,655) dan nomor 21 (0,645). Hanya satu soal yang termasuk kategori sulit, yaitu nomor 26 dengan skor 0,166. Distribusi tingkat kesukaran yang timpang (didominasi soal mudah) berpotensi membuat tes kurang menantang dan tidak mencerminkan variasi kompetensi mahasiswa. Banyak butir soal menunjukkan distractor (pilihan jawaban) yang tidak berfungsi dengan baik, karena terdapat pilihan jawaban yang tidak dipilih sama sekali oleh mahasiswa. Hanya sebagian kecil distraktor yang berfungsi, misalnya soal nomor 2, 9, 11, 20, 23, 26, dan 29 yang masih memerlukan revisi pada pilihan tertentu agar lebih seimbang. Distraktor yang tidak berfungsi menandakan adanya redaksi pengecoh yang terlalu lemah atau terlalu jelas sehingga tidak dipilih mahasiswa.

Hasil penelitian ditemukan bahwa validitas butir soal menunjukkan sejauh mana setiap item mampu mengukur konstruk yang dimaksud (Azwar, 2012). Pada instrumen pertama yang terdiri dari 35 butir soal, hanya sebagian kecil item yang memenuhi kriteria validitas ($r \geq 0,3$). Sebagian besar item memiliki nilai korelasi rendah bahkan negatif yang menandakan adanya soal yang tidak sejalan dengan skor total. Hal ini mengindikasikan bahwa instrumen masih mengandung item yang tidak representatif terhadap kompetensi yang ingin diukur.

Sebaliknya, instrumen kedua memuat 11 butir soal, mayoritas item memiliki nilai korelasi di atas 0,3. Hanya ada dua item yang masih di bawah kriteria, yaitu VAR00019 (0,275) dan VAR00022 (0,247). Dengan demikian, instrumen kedua lebih valid dibandingkan instrumen pertama. Menurut Nunnally & Bernstein (1994), item yang tidak valid sebaiknya dieliminasi atau direvisi agar instrumen lebih fokus mengukur kompetensi yang dituju. Adapun hasil validasi instrumen disajikan pada Tabel 1.

Tabel 1. Hasil Validasi

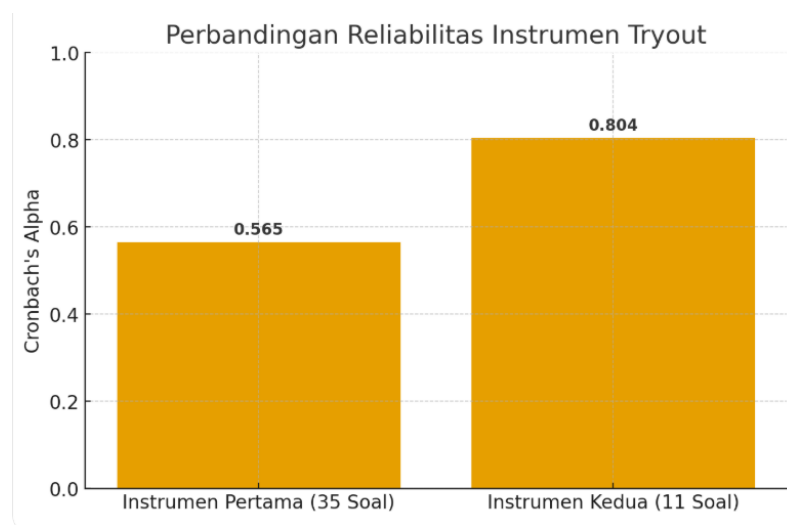
Aspek	Instrumen Pertama	Instrumen Kedua
-------	-------------------	-----------------

Jumlah Butir Soal	35	11
Jumlah Butir Valid	$\pm 7-8$ butir (20–25%)	9 butir (81,8%)
Jumlah Butir Tidak Valid	$\pm 27-28$ butir	2 butir (18,2%)
Cronbach's Alpha	0,565 (rendah–cukup)	0,804 (baik)
Interpretasi	Instrumen kurang layak, perlu perbaikan besar	Instrumen cukup layak, hanya perlu revisi minor

Sedangkan dari hasil reliabilitas ditemukan bahwa reliabilitas mengacu pada konsistensi hasil pengukuran ketika instrumen digunakan dalam kondisi yang sama (Cohen et al., 2018). Hasil uji reliabilitas dengan Cronbach's Alpha menunjukkan perbedaan yang mencolok, yakni: instrumen pertama memperoleh nilai $\alpha = 0,565$, yang menurut George & Mallery (2003) termasuk kategori rendah–cukup, sehingga konsistensinya masih dipertanyakan. Sedangkan instrumen kedua memperoleh nilai $\alpha = 0,804$, yang masuk kategori baik, sehingga reliabilitasnya dapat diterima untuk keperluan evaluasi akademik.

Peningkatan nilai reliabilitas pada instrumen kedua terjadi karena proses seleksi item, di mana butir yang tidak valid dieliminasi sehingga menyisakan soal-soal yang memiliki konsistensi internal lebih tinggi. Hal ini sesuai dengan pandangan Fraenkel & Wallen (2009) yang mengatakan bahwa membuang item dengan korelasi rendah atau negatif dapat meningkatkan reliabilitas instrumen secara keseluruhan. Secara umum hasil uji reliabilitas tes disajikan pada Gambar 1.

Gambar 1 menunjukkan adanya perbandingan reliabilitas instrumen *tryout* pertama dan kedua. Instrumen pertama terdiri dari 35 butir soal dengan nilai Cronbach's Alpha sebesar 0,565, yang termasuk kategori rendah–cukup. Hal ini menunjukkan bahwa konsistensi internal antar butir soal masih lemah, artinya, setiap soal belum saling mendukung dalam mengukur kompetensi yang sama. Sedangkan instrumen kedua terdiri dari 11 butir soal dengan nilai Cronbach's Alpha sebesar 0,804, yang masuk dalam kategori baik. Ini berarti bahwa butir-butir soal dalam instrumen kedua memiliki hubungan yang konsisten dan dapat dipercaya untuk digunakan sebagai alat evaluasi pembelajaran. Secara visual, tinggi batang pada grafik memperlihatkan peningkatan reliabilitas yang signifikan dari 0,565 menjadi 0,804 setelah dilakukan seleksi butir soal. Peningkatan ini menandakan bahwa pengurangan jumlah soal dapat meningkatkan kualitas instrumen, karena butir-butir soal yang tidak valid dan tidak konsisten telah dihapus. Dengan demikian, grafik pada Gambar 1 menggambarkan bahwa kualitas tes tidak ditentukan oleh banyaknya soal, melainkan oleh mutu dan konsistensi tiap butir dalam mengukur konstruk yang sama. Hasil penelitian tersebut menunjukkan bahwa kualitas instrumen evaluasi tidak hanya ditentukan oleh jumlah soal, tetapi lebih pada kualitas item yang digunakan. Instrumen kedua, meskipun jumlah soalnya lebih sedikit tetapi terbukti lebih valid dan reliabel dibandingkan instrumen pertama.



Gambar 1. Hasil Uji Reliabilitas

Hal ini mendukung prinsip bahwa instrumen evaluasi yang baik adalah yang mampu mengukur kompetensi secara tepat dan konsisten (Anastasi & Urbina, 1997). Dengan demikian, instrumen kedua dapat

direkomendasikan untuk digunakan dalam evaluasi pembelajaran untuk mahasiswa PPG Prodi Teknik Kimia Industri. Sementara itu, instrumen pertama memerlukan revisi agar layak digunakan.

Pelaksanaan *tryout* kepada mahasiswa PPG Prodi Teknik Kimia Industri ini dilakukan secara online sehingga mahasiswa dapat dengan mudah mengerjakan soal tetapi bagi administrator sulit mengontrol atau mengawasi secara detail proses pengerjaan mahasiswa. Akibatnya hasil *tryout* mahasiswa cukup bagus karena soal *tryout* yang diberikan kepada mahasiswa tergolong mudah. Sistem *tryout* yang dilakukan secara online kurang memberikan manfaat bagi peserta dalam meningkatkan kualitas persiapan ujian, tetapi juga bagi administrator dalam meningkatkan efektivitas dan akurasi pengelolaan data (Affan & Elhanafi, 2025). Adanya *tryout* ini diharapkan dapat menjadi alat evaluasi terhadap ketercapaian tujuan pembelajaran yang telah dilakukan. Evaluasi merupakan hal penting dalam pembelajaran karena dapat mendorong keterlibatan aktif siswa dalam membentuk pemahamannya (Siregar et al, 2025). Akan tetapi dalam penyusunan soal *tryout* perlu memperhatikan level C1 sampai C6 yang terdapat pada Taksonomi Bloom. Dengan adanya *tryout* untuk mahasiswa PPG dapat melatih kemampuan pedagogi dan berpikir kritis guru. Peningkatan kompetensi guru perlu dilakukan secara berkala untuk meningkatkan keterampilan dalam melaksanakan pengajaran kepada siswa di kelas, khususnya pada mata Pelajaran Teknik Kimia Industri (Nurdini et al., 2024).

SIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa kualitas instrumen *tryout* mahasiswa PPG Teknik Kimia Industri masih perlu ditingkatkan. Dari 35 butir soal yang dianalisis, sebagian besar dinyatakan tidak valid dan hanya sedikit yang memiliki daya beda memadai. Tingkat kesukaran soal didominasi kategori mudah, serta banyak distraktor tidak berfungsi. Reliabilitas instrumen pertama hanya mencapai $\alpha = 0,565$ yang termasuk kategori rendah/cukup. Namun, setelah diseleksi menyisakan 11 soal yang valid sehingga reliabilitas meningkat menjadi $\alpha = 0,804$ dengan kategori baik. Dengan demikian, instrumen kedua dapat direkomendasikan untuk digunakan dalam evaluasi pembelajaran, sedangkan instrumen pertama memerlukan revisi besar-besaran. Hasil penelitian ini menegaskan pentingnya perbaikan kualitas butir soal agar instrumen evaluasi mampu mengukur kompetensi mahasiswa secara lebih tepat dan konsisten. Hasil penelitian ini memberikan kontribusi yang sangat berarti bagi tim pengembang soal PPG karena soal *tryout* yang diberikan kepada mahasiswa terlalu mudah sehingga perlu dilakukan revisi. Keterbatasan penelitian ini adalah menggunakan sampel kecil karena menyesuaikan dengan jumlah mahasiswa PPG yang terdapat di Prodi Teknik Kimia Industri yang ada di Universitas Negeri Malang berjumlah 29 orang sehingga perlu dilakukan penelitian lanjutan dengan sampel yang lebih besar.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Program Pendidikan Profesi Guru (PPG) Universitas Negeri Malang yang telah memberikan pendanaan penelitian dan publikasi artikel ilmiah.

DAFTAR RUJUKAN

- Affan, M. I & Elhanafi, A. M. (2025). Perancangan dan Implementai Aplikasi *Tryout* Online Berbasis Web dengan Fitur Manajemen Soal dan Pendaftaran Terintegrasi. *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, 4(3), 238-245, <https://doi.org/10.69916/jkbt.v4i3.368>
- Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Azwar, S. (2012). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Alexandria, VA: ASCD.
- Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London: Kogan Page.
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8th ed.). London: Routledge.

- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th ed.). Boston: Pearson.
- Fraenkel, J. R., & Wallen, N. E. (2009). *How to Design and Evaluate Research in Education* (7th ed.). New York: McGraw-Hill
- Fullan, M. (2016). *The new meaning of educational change* (5th ed.). Teachers College Press.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Boston: Allyn & Bacon
- Gomez-Coma, L., Díaz-Sainz, G., Fallanza, & Ortiz, A. (2023). Integration of chemical engineering skills in the curriculum of a master course in industrial engineering. *Education for Chemical Engineers*, 45, 68-79
- Hidayatullah, H.A. & Fahmi, A.R. (2025). PBL Berbasis Workshop Sebagai Strategi Pembelajaran Efektif dalam Pendidikan Teknik Kimia. *Didaktika: Jurnal Kependidikan*, 14(2), 2397-2406
- Kemendikbud. (2020). *Merdeka Belajar: Panduan Implementasi Kebijakan Pendidikan Nasional*. Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia
- Kiki, D. R., Saefuddin, & Raharjo, S. (2025). Analisis Profil, Kesiapan, dan Minat Belajar Kimia Siswa Kelas X SMA Negeri 11 Kendari untuk Desain Pembelajaran Berdiferensiasi. *Jurnal Biofiskim: Penelitian dan Pembelajaran IPA*, 7(1), 44-56.
- Nsabayezu, E., Habimana, O., & Nzabalarwa, W. (2025). Examining students' engagement and motivation in organic chemistry through the use of a multimedia-supported flipped classroom approach. *Education for Chemical Engineers*, 53, 102-112
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.
- Nurdini, L., Hendriyani, Shiddiqi, Q. Y. A., Putri, D., Nisrina, H., & Basri, H.M. (2024). Upaya Peningkatan Kompetensi Guru SMK dalam Meningkatkan Mutu Pendidikan Keahlian Teknik Kimia Industri Di Kabupaten Bandung Barat dan Cimahi. *Jurnal Abdimas Kartika Wijayakusuma*, 5(1), 195-202, <https://doi.org/10.26874/jakw.v5i1.380s>
- Rust, C. (2002). The impact of assessment on student learning: How can the research literature practically help to inform the development of departmental assessment strategies and learner-centred assessment practices? *Active Learning in Higher Education*, 3(2), 145–158. <https://doi.org/10.1177/1469787402003002004>
- Siregar, I., Jamril, Adlina, H., Nurjanah, I., Adawiyah, N.R. (2025). Analisis Efektivitas Evaluasi Formatif dan Sumatif dalam Proses Evaluasi Pembelajaran. *KJISE: Kiswah Journal of Islamic Studies an Education*, 1(1).
- Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. San Francisco, CA: Jossey-Bass.
- OECD. (2019). *OECD Future of Education and Skills 2030: OECD Learning Compass 2030*. OECD Publishing.
- Yuriev, E., Burton, M. G., Knott, D., Terrill, A. E., Jackson, N.R.C., & Chen, S. (2025). Integrating Generative AI in Chemistry Education: Enhancing Career-Ready Writing Skills in Pharmaceutical Science, *Journal of Chemical Education*, 102, 1991-2001,