

**Penulisan Butir Soal Mata Pelajaran Bahasa Inggris dengan  
Memanfaatkan Generative Artificial Intelligence (GenAI):  
Pelatihan Peningkatan Kompetensi Guru pada MGMP  
Bahasa Inggris Kecamatan Cipayung Jakarta Timur**

**Fela Dafitri<sup>1\*</sup>, Putri Kamalia Hakim<sup>2</sup>, Fikri Asih Wigati<sup>3</sup>, Nia Pujiawati<sup>4</sup>,  
Maya Rahmawati<sup>5</sup>**

**Universitas Singaperbangsa Karawang**

\*Penulis Korespondensi: [fela.dafitri@fkip.unsika.ac.id](mailto:fela.dafitri@fkip.unsika.ac.id)

**Abstract**

*Assessment in education plays a crucial role in ensuring the quality of both the learning process and student outcomes. One commonly used form of assessment is the written test, presented in the form of test items. However, many teachers face difficulties in writing test items that adhere to proper evaluation principles. With the advancement of technology, Generative Artificial Intelligence (GenAI) has emerged as an innovative solution to support test item development. This community service activity aimed to enhance the competence of English teachers who are members of the Cipayung English MGMP (Subject Teachers' Forum) in East Jakarta in writing test items based on evaluation theory and the use of GenAI. The activity was carried out in three phases: planning, implementation, and evaluation. The training materials covered both theoretical aspects of test writing and hands-on practice using ChatGPT and DeepSeek. The results indicated an improvement in participants' understanding of test item construction and their ability to use GenAI effectively. This activity made a tangible contribution to supporting digital transformation in education, particularly in developing authentic, adaptive, and meaningful assessments.*

**Keywords:** *Test Item Writing, Generative AI, Teacher Training, Digital Assessment*

**Abstrak**

*Penilaian pembelajaran merupakan aspek penting dalam menjamin kualitas proses dan hasil belajar siswa. Salah satu bentuk penilaian yang umum digunakan adalah tes tertulis dalam bentuk butir soal. Namun, banyak guru mengalami kesulitan dalam menulis soal yang sesuai dengan kaidah evaluasi. Seiring perkembangan teknologi, Generative Artificial Intelligence (GenAI) hadir sebagai solusi inovatif dalam mendukung penulisan butir soal. Kegiatan pengabdian kepada masyarakat ini bertujuan untuk meningkatkan kompetensi guru Bahasa Inggris yang tergabung dalam MGMP Cipayung, Jakarta Timur, dalam menulis butir soal berdasarkan teori evaluasi dan pemanfaatan GenAI. Kegiatan dilakukan dalam tiga tahap: perencanaan, pelaksanaan pendampingan, dan evaluasi. Materi pelatihan mencakup teori penulisan soal serta praktik penggunaan ChatGPT dan DeepSeek. Hasil kegiatan menunjukkan peningkatan pemahaman peserta terhadap konsep penulisan soal dan kemampuan mereka dalam menggunakan GenAI secara efektif. Kegiatan ini memberikan kontribusi nyata dalam mendukung transformasi digital pendidikan, khususnya dalam pengembangan asesmen yang autentik, adaptif, dan bermakna.*

**Kata kunci:** *Penulisan Soal, Generative AI, Pelatihan guru, Asesmen Digital*

## **PENDAHULUAN**

Penilaian pembelajaran memiliki peran penting dalam menentukan capaian belajar siswa, serta menjadi dasar pengambilan keputusan dalam proses pembelajaran. Salah satu bentuk penilaian yang umum digunakan adalah tes tertulis dalam bentuk butir soal. Menurut Anderson & Krathwohl (2001), butir soal yang baik harus disusun berdasarkan indikator yang mengacu pada tujuan pembelajaran serta domain kognitif yang ingin dicapai, sehingga dapat mengukur kemampuan siswa secara akurat.

Dalam konteks pengajaran Bahasa Inggris, penulisan butir soal menuntut pemahaman yang mendalam mengenai empat keterampilan berbahasa (listening, speaking, reading, dan writing), serta struktur bahasa dan konteks penggunaannya. Brown & Abeywickrama (2010) menyatakan bahwa penyusunan soal Bahasa Inggris harus memperhatikan keseimbangan antara bentuk (form), fungsi (function), dan tujuan komunikasi agar soal yang dihasilkan mampu merepresentasikan kemampuan berbahasa secara autentik.

Namun, berbagai penelitian menunjukkan bahwa banyak guru mengalami kendala dalam menulis soal yang sesuai dengan prinsip-prinsip evaluasi pembelajaran. Purwanti (2017) menemukan bahwa guru seringkali membuat soal berdasarkan pengalaman pribadi atau sumber tidak resmi tanpa mengacu pada kaidah penulisan soal yang benar. Hal ini berdampak pada rendahnya kualitas soal yang digunakan di sekolah.

Seiring perkembangan teknologi, Generative Artificial Intelligence (GenAI) hadir sebagai alat bantu yang mampu mempercepat dan mempermudah proses penulisan butir

soal. GenAI merupakan bentuk kecerdasan buatan yang dapat menghasilkan teks, gambar, atau konten lainnya secara mandiri berdasarkan permintaan pengguna (prompt). Dwivedi et al. (2023) menjelaskan bahwa GenAI, seperti ChatGPT, memiliki potensi besar dalam dunia pendidikan, terutama dalam pengembangan bahan ajar, soal, dan materi pembelajaran yang adaptif dan personal.

Dalam konteks penilaian, GenAI dapat dimanfaatkan untuk menghasilkan variasi soal dengan cepat, memperkaya tipe soal, dan membantu guru dalam merancang pertanyaan yang sesuai dengan level kognitif tertentu. Kasneci et al. (2023) menekankan bahwa GenAI dapat menjadi partner kolaboratif bagi guru dalam merancang asesmen, selama penggunaannya disertai dengan pemahaman pedagogis yang baik. Selain menghasilkan teks dalam bentuk soal pilihan ganda, uraian, dan materi bacaan, Generative AI seperti ChatGPT juga memiliki kemampuan untuk memproduksi teks multimodal, yaitu menggabungkan elemen verbal dan visual seperti gambar, grafik, atau ilustrasi. Kemampuan ini mendukung pembuatan soal berbasis gambar untuk keterampilan membaca dan menulis, maupun pengembangan bahan ajar visual yang menarik dan kontekstual. Menurut Kohnke et al. (2023), pemanfaatan GenAI dalam pembelajaran bahasa dapat menciptakan pengalaman belajar yang lebih interaktif melalui integrasi teks dan gambar yang relevan dengan dunia nyata. Selain itu, sistem seperti DALL·E yang terintegrasi dengan ChatGPT mampu menghasilkan gambar dari deskripsi teks, sehingga guru dapat menyesuaikan konten

asesmen visual dengan kebutuhan spesifik siswa.

Dalam konteks pembelajaran berdiferensiasi dan Kurikulum Merdeka, fitur ini memungkinkan guru untuk merancang materi dan soal yang adaptif terhadap karakteristik siswa, gaya belajar, serta konteks lokal. Smutny & Schreiberová (2020) menekankan pentingnya fleksibilitas digital dalam merancang bahan ajar yang sesuai dengan kebutuhan pendidikan di berbagai lingkungan sekolah. Oleh karena itu, GenAI tidak hanya berperan sebagai alat bantu otomatisasi, tetapi juga sebagai sarana kolaboratif dalam pengembangan asesmen yang lebih kreatif, personal, dan bermakna bagi peserta didik.

Melihat peluang tersebut, pelatihan ini bertujuan untuk meningkatkan kapasitas guru yang tergabung dalam MGMP Bahasa Inggris Kecamatan Cipayung Jakarta Timur dalam menulis butir soal Bahasa Inggris berbasis teori evaluasi dan pemanfaatan GenAI, sehingga guru tidak hanya memahami kaidah penulisan soal, tetapi juga mampu mengintegrasikan teknologi dalam praktiknya.

## METODE

Dalam kegiatan pengabdian kepada masyarakat ini, pendekatan yang digunakan bertujuan untuk menyelesaikan permasalahan melalui pelatihan yang mencakup aspek teoretis dan praktis. Kegiatan ini dilaksanakan melalui tiga tahap utama, yaitu tahap perencanaan, tahap pelaksanaan pendampingan, dan tahap evaluasi.

1. Tahap Perencanaan
  - a. Melakukan identifikasi terhadap mitra pengabdian, yaitu MGMP Bahasa Inggris Kecamatan Cipayung, termasuk jumlah peserta

serta kebutuhan pelatihan yang diperlukan.

- b. Menyusun materi pelatihan serta merancang prosedur pendampingan yang akan diterapkan.
2. Tahap Pelaksanaan Pendampingan
    - a. Menyampaikan materi mengenai teori penulisan soal tes Bahasa Inggris.
    - b. Menjelaskan konsep Generative AI kepada para guru Bahasa Inggris di MGMP Kecamatan Cipayung.
    - c. Melakukan praktik penggunaan ChatGPT dengan panduan prompt yang sesuai.
  3. Tahap Evaluasi
    - a. Melaksanakan pre-test dan post-test untuk mengukur efektivitas pelatihan.

Fokus utama dalam kegiatan ini adalah pelatihan penulisan butir soal untuk tes Bahasa Inggris serta penerapan AI dalam personalisasi soal. Berdasarkan usulan dari MGMP Bahasa Inggris Kecamatan Cipayung, kegiatan pelatihan ini dilakukan pada hari Senin, 17 Februari 2025. Rundown acara dapat dilihat pada Tabel berikut:

**Tabel 1. Rundown Acara Penulisan Butir Soal Mata Pelajaran Bahasa Inggris dengan Memanfaatkan Generative Artificial Intelligence (GenAI)**

No.	Alokasi Waktu	Materi
1.	13.00 – 13.05	Pembukaan
2.	13.05 – 13.10	Sambutan Kepala SMP 230 Jakarta
3.	13.10 – 13.15	Sambutan Ketua MGMP Bahasa Inggris Cipayung Jakarta Timur
4.	13.15 – 14.15	Materi I: <i>Item Writing for Language Test</i>
5.	14.15 – 15.15	Materi II: <i>Personalizing Test Items with AI</i>
6.	15.15 – 15.30	Penutupan

## HASIL DAN PEMBAHASAN

### 1. Tahap Perencanaan

Tahap perencanaan merupakan fondasi penting dalam pelaksanaan kegiatan pengabdian kepada masyarakat ini, karena pada tahap inilah kebutuhan mitra dan strategi pelaksanaan kegiatan diidentifikasi dan dirumuskan secara sistematis. Langkah pertama yang dilakukan oleh tim pelaksana adalah menjalin komunikasi dan melakukan pertemuan langsung dengan Ketua Musyawarah Guru Mata Pelajaran (MGMP) Bahasa Inggris Kecamatan Cipayung. Tujuan utama dari pertemuan ini adalah untuk menggali secara mendalam kebutuhan para guru Bahasa Inggris yang tergabung dalam MGMP, serta menyesuaikan bentuk kegiatan pelatihan dengan konteks nyata yang mereka hadapi di lapangan.

Melalui diskusi yang konstruktif, diperoleh informasi bahwa para guru sangat membutuhkan pelatihan yang dapat mendukung mereka dalam menyusun soal untuk Ujian Akhir Sekolah (UAS), terutama soal-soal Bahasa Inggris yang menuntut keterampilan berpikir tingkat tinggi dan autentisitas konteks. Berdasarkan kebutuhan tersebut, disepakati tanggal pelaksanaan pelatihan dan dipilih dua fokus materi utama, yaitu *Item Writing for Language Test* dan *Personalization of Test Items Using Generative AI*. Kedua materi ini dirancang tidak hanya untuk meningkatkan kompetensi teknis guru dalam menulis soal, tetapi juga untuk memperkenalkan pendekatan inovatif berbasis teknologi, yang relevan dengan perkembangan dunia pendidikan saat ini.

Setelah agenda pelatihan ditetapkan, narasumber menyusun materi pelatihan secara sistematis dan merancang prosedur pendampingan

yang terstruktur. Materi pertama, *Item Writing for Language Test*, dirancang untuk merevitalisasi pemahaman para guru terhadap prinsip-prinsip dasar dalam penulisan butir soal, seperti kesesuaian indikator soal, kejelasan bahasa, kevalidan konten, dan keterukuran jawaban. Kegiatan ini penting untuk memastikan bahwa soal-soal yang disusun mampu mengukur capaian kompetensi siswa secara akurat.

Materi kedua, yaitu *Personalization of Test Items Using Generative AI*, difokuskan pada pengenalan dan pemanfaatan teknologi Generative AI, khususnya ChatGPT, dalam proses penulisan soal. Para guru diperkenalkan pada cara menggunakan GenAI untuk menghasilkan teks bacaan yang sesuai dengan kebutuhan pembelajaran di sekolah masing-masing, menyusun soal berdasarkan teks tersebut, serta memanfaatkan fitur visualisasi dengan membuat gambar yang relevan dengan konten soal. Selain itu, guru juga dibimbing untuk menyusun *prompt* secara tepat agar hasil dari GenAI sesuai dengan harapan dan kebutuhan konteks pembelajaran. Dalam hal ini, guru dilatih untuk mengikuti empat prinsip utama dalam menyusun *prompt*, yaitu: (1) *Identify your role* (menentukan peran yang diharapkan dari GenAI), (2) *Provide context* (memberikan latar belakang atau konteks soal), (3) *Be specific and clear* (menggunakan bahasa yang jelas dan rinci), dan (4) *Set parameters* (menentukan batasan atau kriteria hasil yang diinginkan).

Perencanaan yang matang ini menjadi langkah strategis dalam memastikan bahwa kegiatan pengabdian tidak hanya bersifat seremonial, tetapi memberikan dampak nyata terhadap peningkatan kapasitas guru, baik dalam aspek pedagogis

maupun teknologi. Dengan demikian, diharapkan kegiatan ini mampu menjawab tantangan pembelajaran abad ke-21 dan mendukung transformasi digital dalam pendidikan.

## 2. Tahap Pelaksanaan Pendampingan

Tahap pelaksanaan pendampingan merupakan momen kunci dalam kegiatan pengabdian kepada masyarakat ini, di mana transfer pengetahuan dan keterampilan secara langsung dilakukan kepada para guru sebagai peserta pelatihan. Kegiatan pelatihan penulisan butir soal Bahasa Inggris berbasis *Generative Artificial Intelligence* (GenAI) dilaksanakan secara tatap muka dan diikuti oleh 15 orang guru Bahasa Inggris yang merupakan anggota aktif MGMP Kecamatan Cipayung, Jakarta Timur. Kehadiran para peserta yang tepat waktu serta keterlibatan aktif selama kegiatan menunjukkan komitmen dan antusiasme yang tinggi terhadap pengembangan profesionalisme mereka.



Gambar 1. Foto Bersama Narasumber dengan Peserta

Pelatihan dilaksanakan sesuai dengan alur kegiatan yang telah dirancang dalam *rundown*, dan dibagi menjadi dua sesi utama, masing-masing dengan fokus yang berbeda. Pada sesi pertama, materi yang disampaikan adalah *Item Writing for*

*Language Test*. Sesi ini bertujuan untuk menguatkan kembali pemahaman para guru mengenai prinsip-prinsip dasar penyusunan butir soal Bahasa Inggris yang baik, mencakup aspek validitas, reliabilitas, kejelasan bahasa, dan kesesuaian dengan indikator pembelajaran. Narasumber juga menekankan pentingnya keseimbangan antara aspek kognitif dan kontekstual dalam penulisan soal, agar soal tidak hanya mengukur hafalan, tetapi juga kemampuan berpikir kritis siswa.



Gambar 2. Penyampaian Materi I (*Item Writing for Language Test*)

Meskipun materi disampaikan secara interaktif, tanggapan dari peserta pada sesi ini relatif terbatas. Hal ini dapat dipahami karena sebagian besar peserta adalah guru dengan pengalaman mengajar lebih dari sepuluh tahun, dan telah terbiasa menyusun soal dalam berbagai format ujian. Salah satu guru bahkan mengemukakan pandangannya bahwa teori tentang pengecoh atau *distractor* dalam soal pilihan ganda mulai kehilangan relevansinya dalam konteks penilaian di sekolah, yang kini lebih mengarah pada soal berbasis teks dan penalaran.

Sesi kedua difokuskan pada eksplorasi penggunaan teknologi GenAI dalam penulisan soal. Narasumber memperkenalkan konsep dasar *Generative Artificial Intelligence*, serta memperlihatkan

fungsinya dalam menghasilkan bahan ajar dan soal secara adaptif. Dua platform AI yang digunakan dalam pelatihan ini adalah ChatGPT dan DeepSeek. Keduanya dipilih agar peserta dapat membandingkan hasil keluaran dari masing-masing sistem, baik dari segi kualitas bahasa, kesesuaian konten, maupun kreativitas dalam menghasilkan variasi soal.



**Gambar 3. Penyampaian Materi II  
(*Personalizing Test Items with AI*)**

Dalam sesi praktik, peserta diajak untuk langsung menggunakan kedua platform dengan bimbingan narasumber. Mereka diberi kesempatan untuk merancang *prompt* yang efektif, sesuai dengan empat prinsip utama: (1) *Identify your role*, (2) *Provide context*, (3) *Be specific and clear*, dan (4) *Set parameters*. Dengan menggunakan prinsip ini, peserta dapat meminta AI untuk menghasilkan teks bacaan, membuat gambar ilustratif, dan menyusun soal berdasarkan kebutuhan spesifik konteks sekolah masing-masing.

Respons peserta pada sesi ini sangat positif. Banyak di antara mereka yang baru pertama kali menggunakan AI untuk tujuan pendidikan, sehingga pelatihan ini membuka wawasan baru mengenai potensi teknologi dalam menunjang tugas-tugas guru,

khususnya dalam pembuatan soal yang kreatif, kontekstual, dan hemat waktu. Beberapa guru juga menyampaikan bahwa pelatihan ini sangat relevan dengan tantangan mereka saat ini, terutama dalam menyusun soal yang sesuai dengan kemampuan siswa, sekaligus menarik secara visual dan kognitif.

Secara keseluruhan, tahap pelaksanaan pendampingan ini berhasil menciptakan suasana belajar yang kolaboratif, mempertemukan pengalaman praktis para guru dengan pendekatan inovatif berbasis teknologi. Kegiatan ini tidak hanya meningkatkan kapasitas teknis guru dalam menyusun soal, tetapi juga memperluas cakrawala mereka dalam memanfaatkan AI sebagai mitra dalam proses pengembangan materi ajar yang efektif dan menarik.

### 3. Tahap Evaluasi

Tahap evaluasi dalam kegiatan pengabdian ini dilakukan dalam dua bentuk, yakni *pre-test* dan *post-test*, yang bertujuan untuk mengukur pengetahuan awal peserta serta perkembangan pemahaman mereka setelah mengikuti pelatihan. Evaluasi ini penting dilakukan sebagai indikator efektivitas pelatihan serta sebagai dasar perencanaan pelatihan lanjutan di masa depan.

#### Pre-Test

Pre-test dilakukan pada awal kegiatan dengan menggunakan aplikasi *Slido* yang memungkinkan peserta menjawab secara daring dan anonim. Terdapat tiga pertanyaan utama yang diajukan untuk mengukur tingkat pemahaman dan pengalaman peserta terkait penggunaan *Generative Artificial Intelligence* (GenAI) dalam penulisan soal.

- Pertanyaan pertama menanyakan apakah peserta pernah menggunakan AI dalam menyusun soal ujian. Hasilnya menunjukkan bahwa 56% dari peserta menyatakan belum pernah menggunakan AI dalam konteks tersebut.
- Pertanyaan kedua menyoroti jenis aplikasi AI yang pernah digunakan dalam pembuatan soal. Jawaban yang muncul antara lain *Wordwall*, *Meta AI*, dan *ChatGPT*, meskipun sebagian besar peserta kembali menjawab bahwa mereka belum pernah menggunakan aplikasi AI untuk tujuan tersebut.
- Pertanyaan ketiga bertujuan untuk menggali pemahaman peserta tentang fungsi AI dalam mendukung pembuatan soal, khususnya teks bacaan. Beberapa peserta menjawab bahwa AI digunakan untuk mencari contoh teks, menyusun soal latihan, atau sebagai referensi semata. Namun, sebagian besar responden tetap menjawab belum pernah menggunakan AI untuk fungsi-fungsi tersebut.

Berdasarkan hasil pre-test ini, dapat disimpulkan bahwa sebagian besar peserta memiliki pengetahuan dan pengalaman yang masih terbatas dalam memanfaatkan teknologi AI untuk kepentingan penyusunan soal Bahasa Inggris. Hal ini menunjukkan bahwa pelatihan yang diberikan sangat relevan dan sesuai dengan kebutuhan mereka.



Gambar 4. Penggunaan Barcode untuk Pre-Test

### Post-Test

Post-test dilakukan secara lisan melalui sesi tanya jawab di akhir kegiatan. Evaluasi ini bersifat reflektif, di mana peserta diminta untuk menyampaikan pemahaman dan kesan mereka setelah mengikuti seluruh rangkaian pelatihan. Berdasarkan hasil diskusi, seluruh peserta (100%) menyatakan bahwa mereka menjadi lebih memahami konsep serta cara kerja GenAI dalam menyusun butir soal, khususnya dalam konteks pembelajaran Bahasa Inggris.

Selain menunjukkan pemahaman yang meningkat, para peserta juga mengungkapkan antusiasme yang tinggi terhadap pemanfaatan AI dalam pekerjaan mereka. Beberapa guru bahkan menyatakan ketertarikannya untuk langsung mencoba menyusun soal ujian sekolah menggunakan bantuan AI, karena dianggap dapat menghemat waktu, memberikan variasi soal yang lebih kaya, serta membantu menghasilkan materi yang lebih kontekstual dan menarik bagi siswa.

Secara keseluruhan, tahap evaluasi ini menunjukkan adanya peningkatan signifikan dalam aspek pengetahuan dan sikap peserta terhadap penggunaan teknologi AI dalam dunia pendidikan, khususnya dalam praktik

penilaian. Hal ini memperkuat relevansi dan urgensi pelatihan semacam ini untuk diterapkan secara lebih luas di masa mendatang.

### KESIMPULAN

Kegiatan pelatihan penulisan butir soal Bahasa Inggris berbasis Generative Artificial Intelligence (GenAI) yang dilaksanakan bagi guru-guru MGMP Bahasa Inggris Kecamatan Cipayung berhasil mencapai tujuannya dalam meningkatkan pemahaman dan keterampilan peserta. Melalui tahapan perencanaan yang matang, pelaksanaan yang kolaboratif, dan evaluasi yang terstruktur, kegiatan ini memberikan kontribusi nyata terhadap peningkatan kapasitas guru dalam merancang soal yang sesuai dengan prinsip evaluasi pembelajaran sekaligus memanfaatkan teknologi terkini. Hasil pre-test menunjukkan bahwa mayoritas peserta belum memiliki pengalaman dalam menggunakan AI untuk keperluan penilaian, sedangkan post-test menunjukkan perubahan signifikan dalam pemahaman serta minat peserta untuk mengintegrasikan GenAI dalam praktik profesional mereka.

Pelatihan ini juga membuktikan bahwa penggunaan GenAI, seperti ChatGPT dan DeepSeek, mampu memperkaya proses penulisan soal dengan menyediakan teks, ilustrasi, dan pertanyaan yang relevan dan kontekstual. Selain itu, pelatihan ini memperkenalkan pendekatan prompt engineering sebagai keterampilan penting dalam mengarahkan kinerja AI untuk menghasilkan output yang sesuai dengan kebutuhan pendidikan. Ke depan, pelatihan serupa dapat diperluas ke MGMP lain dan disesuaikan dengan kurikulum nasional, seperti Kurikulum Merdeka, agar pemanfaatan teknologi tidak hanya meningkatkan efisiensi

kerja guru, tetapi juga mendukung pembelajaran yang lebih adaptif, kreatif, dan bermakna.

### REFERENSI

- Kohnke, L., Zou, D., & Zhang, R. (2023). Exploring AI-generated content in English language learning: Benefits, challenges, and ethical implications. *Computer Assisted Language Learning*, 1–24. <https://doi.org/10.1080/09588221.2023.2187545>
- Smutny, P., & Schreiberová, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Education Sciences*, 10(10), 269. <https://doi.org/10.3390/educsci10100269>
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson Education.
- Dwivedi, Y. K., Hughes, D. L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., & Al-Debei, M. M. (2023). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 68, 102585.

- <https://doi.org/10.1016/j.ijinfomgt.2022.102585>
- Kasneci, E., Sessler, K., Bittner, E. M., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274.  
<https://doi.org/10.1016/j.lindif.2023.102274>
- Miller, M. D., Linn, R. L., & Gronlund, N. E. (2013). *Measurement and assessment in teaching* (11th ed.). Pearson Education.
- Puspendik. (2019). *Panduan penulisan soal*. Pusat Penilaian Pendidikan, Kementerian Pendidikan dan Kebudayaan.
- Purwanti, E. (2017). Teachers' ability in constructing English test items. *Journal of English Language Teaching and Linguistics, 2*(3), 213–222.  
<https://doi.org/10.21462/jeltl.v2i3.81>