

BIBLIOGRAPHY

- Agostini, D., & Picasso, F. (2024). Large language models for sustainable assessment and feedback in higher education. *Intelligenza Artificiale*, 18(1), 121–138. <https://doi.org/10.3233/IA-240033>
- Ali, M. K., Ali, A. M., & 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. <https://doi.org/10.54373/ijset.v4i1.1649>
- Anisah, G. (2021). Kerangka Konsep Assessment of Learning, Assessment for Learning, dan Assessment as Learning serta Penerapannya pada Pembelajaran. *Jurnal Pendidikan dan Kajian Keislaman*, 65–76.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Branch, R. M. (2009). Instructional Design: The ADDIE Approach. In *Instructional Design: The ADDIE Approach*. Springer US. <https://doi.org/10.1007/978-0-387-09506-6>
- Chen, B., Zhang, Z., Langrené, N., & Zhu, S. (2025). Unleashing the potential of prompt engineering for large language models. *Patterns*, 6(6), 101260. <https://doi.org/10.1016/j.patter.2025.101260>
- Darmayasa, P. R., Bagus, I., Pascima, N., & Agustini, K. (2025). Pengembangan Sistem Penilaian Keaktifan Belajar Otomatis Berbasis Learning Management System

(LMS) Dengan Retrieval-Augmented Generation (RAG). *Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika (KARMAPATI)*, 14(3).

Ehsan, Md. A., Hasan, A. S. M. M., Shahnour, K. B., & Tasneem, S. S. (2025). *Automatic Question & Answer Generation Using Generative Large Language Model (LLM)*. <http://arxiv.org/abs/2508.19475>

Flores Limo, F., Tiza, D., Roque, M., Herrera, E., Muñoz Murillo, J., Jinchuña Huallpa, J., Ariza Flores, V., Castillo, A., Peña, P., & Carranza, C. (2023). Personalized tutoring: ChatGPT as a virtual tutor for personalized learning experiences. *Przestrzen Społeczna*, 23, 293–312.

Handayani, R., & Wulandari, D. (2021). Modern Assessment dalam Menyongsong Pembelajaran Abad 21 dan Hambatan di Negara Berkembang. *Jurnal Pendidikan Edutama*, 8(1), 13. <https://doi.org/10.30734/jpe.v8i1.1363>

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. <https://doi.org/10.1186/s40561-024-00310-z>

Jiang, B., Li, X., Yang, S., Kong, Y., Cheng, W., Hao, C., & Lin, Q. (2022). Data-Driven Personalized Learning Path Planning Based on Cognitive Diagnostic Assessments in MOOCs. *Applied Sciences*, 12(8), 3982. <https://doi.org/10.3390/app12083982>

Juryatina, J., Haris, A., Uyun, N., & Ekahidayatullah, M. (2025). Evaluasi Praktik Asesmen Diagnostik IPA Untuk Mengidentifikasi Kesulitan Belajar Siswa Sekolah Dasar. *JUPEIS: Jurnal Pendidikan Dan Ilmu Sosial*, 4(3), 547–555. <https://doi.org/10.57218/jupeis.Vol4.Iss3.1722>

Khairunnisa, K., Ningsih, K., & Panjaitan, R. G. P. (2022). Kelayakan media booklet pada materi interaksi makhluk hidup dan lingkungannya kelas VII. *Jurnal Pendidikan Informatika Dan Sains*, *11*(2), 148–155. <https://doi.org/10.31571/saintek.v11i2.4095>

Leighton, J. P., & Gierl, M. J. (2007). Why cognitive diagnostic assessment? In *Cognitive Diagnostic Assessment for Education: Theory and Applications* (pp. 3–18). Cambridge University Press. <https://doi.org/10.1017/CBO9780511611186.001>

Maemonah, M. (2018). *Asesmen Pembelajaran* (F. Yuliawati, Ed.; 1st ed.). PGMI PRESS UIN SUKA. <http://www.pgmi.uin-suka.ac>

Maulida, U. (2022). PENGEMBANGAN MODUL AJAR BERBASIS KURIKULUM MERDEKA. *Tarbawi : Jurnal Pemikiran Dan Pendidikan Islam*, *5*(2), 130–138. <https://doi.org/10.51476/tarbawi.v5i2.392>

Maut, W. O. A. (2022). Asesmen Diagnostik dalam Implementasi Kurikulum Merdeka (IKM) di SD Negeri 1 Tongkuno Kecamatan Tongkuno Kabupaten Muna Sulawesi Tenggara. *Http://Ejurnal.Pps.Ung.Ac.Id/Index.Php/Dikmas*, *02*(4), 1305–1312. <https://doi.org/http://dx.doi.org/10.37905/dikmas.2.4.1305-1312.2022>

Mushthofa, Wahyono, Asfarian, A., Ramadhan, D. A., Putro, P. P., Wisnubhadra, I., Saputra, B., & Pratiwi, H. (2021). *INFORMATIKA SMA KELAS X* (C. Tulalessy & C. E. Nugraheni, Eds.). Pusat Kurikulum dan Perbukuan Balitbang, Kemendikd., 2021.

Nazilah, A. S. (2024). Diagnostic Assessment in Differentiated Learning: Supporting

Learners' Needs for Improved Learning Achievement. *Research Journal on Teacher Professional Development*, 2(1), 87–92.
<https://doi.org/10.21580/rjtpd.v2i1.20662>

Ningsi, P. A., Purwaningsi, S., & Darmaji. (2021). *Pengembangan Penuntun Praktikum Elektronik Berbasis Keterampilan Proses Sains Materi Suhu dan Kalor untuk SMP/Mts*. <https://doi.org/https://doi.org/10.5281/zenodo.4734431>

Nouriska, S., Untoro, M. C., Afriansyah, A., Praseptiawan, M., Yulita, W., & Ashari, I. F. (2023). USER EXPERIENCE ANSWER SYSTEM AUTOMATICALLY WITH USER CENTERED DESIGN AND USER EXPERIENCE QUESTIONNAIRE-SHORT. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 9(1), 81–88. <https://doi.org/10.33480/jitk.v9i1.4152>

Nurhasanah, N., Azhari, A., Berutu, K., Putra, T. J., Hasibuan, R. H., & Nasution, I. (2023). Evaluasi Pembelajaran Dikelas. *Jurnal Motivasi Pendidikan Dan Bahasa*, 1(2), 257–270. <https://doi.org/10.59581/jmpb-widyakarya.v1i2.485>

Nurhuda, T., Budiarta, N. E., Nyoman, I., Wiradika, I., Gede, I., Subawa, B., Ketut, I., Pradnyana, A., & Bali, P. B. (2025). Rancang Bangun Sistem Deteksi kemampuan awal menulis saintifik anak berbasis LLM. *Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika (KARMAPATI)*, 14(3).

Pavlova, M. (2009). *Technology and Vocational Education for Sustainable Development* (M. Pavlova, Ed.; Vol. 10). Springer Netherlands. <https://doi.org/10.1007/978-1-4020-5279-8>

Putra, I. N. T. A., Driya, P. D., Atmaja, I. G. B. W., Mahendra, K., Supriatmaja, G. A.,

- & Pratama, I. P. M. Y. (2025). REKAYASA SISTEM INFORMASI DIEVALUASI BERBASIS MODEL WATERFALL: EKSPERIMEN KEANDALAN BLACK-BOX DAN OPTIMALISASI UX MELALUI UEQ. *Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika (KARMAPATI)*, 14(2).
- Qurtubi, A., Hakim, M., & Afandi, B. (2024). Pengembangan Instrumen Asesmen Diagnostik Kognitif Berbasis Android dan Desktop pada Pembelajaran IPA Kelas VIII SMP Daruttaqwa Jombang. *Jurnal Pendidikan Biologi Undiksha*, 10(3), 71–84. <https://doi.org/10.23887/jjpb.v10i3.68132>
- Rangkuti, A. N. (2016). *METODE PENELITIAN PENDIDIKAN (Pendekatan Kuantitatif, Kualitatif, PTK, dan Penelitian Pengembangan)* (M. S. Lubis, Ed.; 1st ed.). Citapustaka Media.
- Riyanto, A. (2026). The Effectiveness of Differentiated Learning in Indonesian Language Subjects at Elementary Schools. *JOURNAL OF INNOVATION AND RESEARCH IN PRIMARY EDUCATION*, 5(1), 312–324. <https://doi.org/https://doi.org/10.56916/jirpe.v5i1.2883>
- Sabri, S. (2013). Item analysis of student comprehensive test for research in teaching beginner string ensemble using model based teaching among music students in public universities. *International Journal of Education and Research*, 1, 1–14. www.ijern.com
- Scaria, N., Dharani Chenna, S., & Subramani, D. (2024). Automated Educational Question Generation at Different Bloom's Skill Levels Using Large Language

Models: Strategies and Evaluation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): 14830 LNAI* (pp. 165–179). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-031-64299-9_12

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6), 103–108. <https://doi.org/10.9781/ijimai.2017.09.001>

Shin, J., Chen, F., Lu, C., & Bulut, O. (2022). Analyzing students' performance in computerized formative assessments to optimize teachers' test administration decisions using deep learning frameworks. *Journal of Computers in Education*, 9(1), 71–91. <https://doi.org/10.1007/s40692-021-00196-7>

Sideridis, G., Ghamdi, H., & Zamil, O. (2023). Contrasting multistage and computer-based testing: score accuracy and aberrant responding. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1288177>

Sugihartini, N., & Yudiana, K. (2018). ADDIE SEBAGAI MODEL PENGEMBANGAN MEDIA INSTRUKSIONAL EDUKATIF (MIE) MATA KULIAH KURIKULUM DAN PENGAJARAN. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 15(2), 277. <https://ejournal.undiksha.ac.id/index.php/JPTK/issue/view/851>

Sugiyono. (2013). *METODE PENELITIAN KUANTITATIF, KUALITATIF, DAN R&D* (19th ed.). ALFABETA. www.cvalfabet.com

Sulistianingsih, & Agus Wismanto. (2024). EFEKTIVITAS ASESMEN DIAGNOSTIK DALAM PEMBELAJARAN BERDIFERENSIASI DENGAN PENDEKATAN TEACHING AT THE RIGHT LEVEL (TARL) DI SMA. *Jurnal Bastra (Bahasa Dan Sastra)*, 9(3), 664–675. <https://doi.org/10.36709/bastra.v9i3.544>

Sumarsono, S. (2022). Perancangan Fitur Learning Management System (LMS) untuk Penanaman Nilai Berbasis Cognitive Moral Development. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 11(3), 301–309. <https://doi.org/10.23887/janapati.v11i3.54106>

Susandi, A. D., Zulaiha, F., Sudirman, S., Wahyuningrum, E., Yumiati, Y., & Erni, E. (2025). Pelatihan Pembuatan Modul Ajar Berdasarkan Hasil Tes Diagnostik dan Penggunaan Aplikasi PMM Bagi Guru Sekolah Dasar di Kecamatan Plumbon. *Jurnal SOLMA*, 14(1), 1064–1074. <https://doi.org/10.22236/solma.v14i1.16761>

Tang, F., & Zhan, P. (2021). Does Diagnostic Feedback Promote Learning? Evidence From a Longitudinal Cognitive Diagnostic Assessment. *AERA Open*, 7. <https://doi.org/10.1177/23328584211060804>

Tobler, S. (2024). Smart grading: A generative AI-based tool for knowledge-grounded answer evaluation in educational assessments. *MethodsX*, 12, 102531. <https://doi.org/10.1016/j.mex.2023.102531>

Toprak, T. E., & Cakir, A. (2021). Examining the L2 reading comprehension ability of adult ELLs: Developing a diagnostic test within the cognitive diagnostic assessment framework. *Language Testing*, 38(1), 106–131. <https://doi.org/10.1177/0265532220941470>

Von Davier, A. A., Mislevy, R. J., & Hao, J. (2021). *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment* (A. A. von Davier, R. J. Mislevy, & J. Hao, Eds.). Springer International Publishing. <https://doi.org/10.1007/978-3-030-74394-9>

Wahyuni, D. S., Sudarmada, N., Putu Era Marsakawati, N., Gede Bendesa Subawa, I., Ketut Andika Pradnyana, I., Ariadi, G., & Surjo Abdi, A. (2025). *PENGEMBANGAN MEDIA AJAR MULTIMODAL DI SDN 2 NGIS KARANGASEM*.

Wanichsan, D., Panjaburee, P., & Chookaew, S. (2021). Enhancing knowledge integration from multiple experts to guiding personalized learning paths for testing and diagnostic systems. *Computers and Education: Artificial Intelligence*, 2, 100013. <https://doi.org/10.1016/j.caeai.2021.100013>

Weissman, A., Belov, D. I., & Armstrong, R. (2007). *Information-Based Versus Number-Correct Routing in Multistage Classification Tests*. <https://www.researchgate.net/publication/237288650>

Wicaksono, K. A. D., Handayanto, A., & Happy, N. (2020). Pengembangan E-Modul Matematika Berbasis Pendekatan Kontekstual Berbantu Media Powerpoint untuk Meningkatkan Pemahaman Konsep Matematika Siswa pada Materi Program Linear. *Imajiner: Jurnal Matematika Dan Pendidikan Matematika*, 2(6), 461–466. <https://doi.org/10.26877/imajiner.v2i6.6668>

Wiradika, I. N. I. (2024). *Asesmen dan Evaluasi Pendidikan Dengan Program R* (I. M. G. Sunarya, Ed.). Jalan Udayana, Nomor 11 Singaraja, Bali Telp/Faks. (0362)

26327. www.penerbit.undiksha.ac.id

Yamamoto, K., Shin, H. J., & Khorramdel, L. (2018). Multistage Adaptive Testing Design in International Large-Scale Assessments. *Educational Measurement: Issues and Practice*, 37(4), 16–27. <https://doi.org/10.1111/emip.12226>

Yigci, D., Eryilmaz, M., Yetisen, A. K., Tasoglu, S., & Ozcan, A. (2025). Large Language Model-Based Chatbots in Higher Education. *Advanced Intelligent Systems*, 7(3). <https://doi.org/10.1002/aisy.202400429>

Zhang, Z., Hamzah, M. I., & Jamaludin, K. A. (2025). A Framework for Constructing Cognitive Diagnostic Reports on English Reading for Chinese University Students. *Forum for Linguistic Studies*, 7(3), 643–655. <https://doi.org/10.30564/fls.v7i3.8674>

Zheng, Y., & Chang, H.-H. (2015). On-the-Fly Assembled Multistage Adaptive Testing. *Applied Psychological Measurement*, 39(2), 104–118. <https://doi.org/10.1177/0146621614544519>

