

DAFTAR PUSTAKA

- Aini, N., Sari, A., Antia, V., Daimah, U. S., Muhakimah, I., & Dewanti, S. S. (2024). Konstruksi instrumen tes kemampuan pemecahan masalah menggunakan teori respon butir. *Teorema: Teori Dan Riset Matematika*, 09((02)), 193–206. <https://doi.org/DOI:http://dx.doi.org/10.25157/teorema.v9i2.14867>
- Aji, N. W., & Retnawati, H. (2024). Perbandingan kecocokan model analisis teori respons butir pada data politomi. *Jurnal Statistika Dan Sains Data*, 1(2), 31–47. <https://doi.org/10.21831/jssd.v1i2.18484>
- Akdon, & Riduwan. (2009). *Rumus dan Data dalam Analisis Statistik untuk Penelitian*. Bandung: Alfabeta.
- Alagoz, C. (2000). *Scoring tests with dichotomous and polytomous items*. Gazi University
- Amanda, N., & Nusantara, T. (2021). Analisis berpikir kritis siswa terhadap pemecahan masalah matematika di MTs Surya Buana Malang. *Jurnal Pendidikan Matematika Dan Sains*, 8(2), 89–92. <https://doi.org/10.21831/jpms.v8i2.19660>
- Ambarwati, R., & Alim, F. (2023). An analysis of students' critical thinking ability in solving jumping task questions based on learning style. *Journal of Innovative Mathematics Learning*, 6(4), 319–327. <https://doi.org/10.22460/jiml.v6i4.p21727>
- Amelia, R. N., & Kriswantoro, K. (2017). Implementation of item response theory for analysis of test items quality and students' ability in chemistry. *JKPK (Jurnal Kimia Dan Pendidikan Kimia)*, 2(1), 1. <https://doi.org/10.20961/jkpk.v2i1.8512>
- An Nabil, N. R., Wulandari, I., Yamtinah, S., Ariani, S. R. D., & Ulfa, M. (2022). Analisis indeks aiken untuk mengetahui validitas isi instrumen asesmen kompetensi minimum berbasis konteks sains kimia. *PAEDAGOGIA*, 25(2), 184. <https://doi.org/10.20961/paedagogia.v25i2.64566>
- Angraini, L., & Wahyuni, A. (2021). The effect of concept attainment model on mathematical critical thinking ability. *International Journal of Instruction*, 14, 727–742. <https://doi.org/10.29333/iji.2021.14144a>
- Arikunto, S. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arini, R., Riansi, E. S., Fadillah, I., & Hasan, M. K. (2024). Peran tes objektif dan subjektif dalam mengukur keterampilan berfikir matematis siswa sekolah dasar. *Jurnal Inovasi Dan Riset Pendidikan Matematika*, 5(3), 236–244.

- Bahar, R., & Retnawati, H. (2022). Analisis karakteristik soal kemampuan koneksi matematika penskoran politimus. *Jurnal Tarbiyah*, 29, 195. <https://doi.org/10.30829/tar.v29i2.1650>
- Baker, F., & Kim, S.-H. (2017). *The Basics of Item Response Theory Using R*. Springer. <https://doi.org/10.1007/978-3-319-54205-8>
- Brookhart, S. M. (2010). *How to Assess Higher-Order Thinking Skills in Your Classroom*. ASCD. www.ascd.org/memberbooks
- Ciptari, P. A., Purwanti, N. K. R., & Erawati, N. K. (2024). Comparison of the classical test theory and the rasch model in the analysis of mastery test of concept systems of linear (SPLDV). *Prima: Jurnal Pendidikan Matematika*, 8(2), 390. <https://doi.org/10.31000/prima.v8i2.8647>
- Dai, S., Vo, T., Kehinde, O., He, H., Xue, Y., Demir, C., & Wang, X. (2021). Performance of polytomous irt models with rating scale data: an investigation over sample size, instrument length, and missing data. *Frontiers in Education*, 6. <https://doi.org/10.3389/educ.2021.721963>
- Demars. (2010). Item response theory : understanding statistics measurement. *Oxford University Press*. <https://doi.org/https://doi.org/10.15713/ins.mmj.3>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers. <https://doi.org/10.4324/9781410605269>
- Ennis, R. H. (1996). *Critical Thinking Dispositions: Their Nature and Assessability*. Resnick.
- Ennis, R. H. (2011). *The nature of critical thinking: an outline of critical thinking dispositions and abilities*. https://education.illinois.edu/docs/default-source/faculty-documents/robert-ennis/thenatureofcriticalthinking_51711_000.pdf
- Ennis, R. H. (2018). Critical thinking across the curriculum: a vision. *Topoi*, 37(1), 165–184. <https://doi.org/10.1007/s11245-016-9401-4>
- Faiz, A., Putra, N. P., & Fajar, N. (2022). Memahami makna tes, pengukuran (measurement), penilaian (assessment), dan evaluasi (evaluation). *Jurnal Education and Development Institut Pendidikan Tapanuli Selatan*, 10(3). <https://doi.org/10.31004/obsesi.v5i2.972>
- Fajrianthi, F., Hendriani, W., & Septarini, B. G. (2016). Pengembangan tes berpikir kritis dengan pendekatan item response theory. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(1), 45–55. <https://doi.org/10.21831/pep.v20i1.6304>

- Falani, I., Akbar, M., & Naga, D. S. (2020a). Comparison of the accuracy of item response theory models in estimating student's ability. *Journal of Educational Science and Technology*, 6(2), 178–184. <https://doi.org/https://doi.org/10.26858/est.v6i2.13295>
- Falani, I., Akbar, M., & Naga, D. S. (2020b). The precision of students' ability estimation on combinations of item response theory models. *International Journal of Instruction*, 13(4), 545–558. <https://doi.org/https://doi.org/10.29333/iji.2020.13434a>
- Falani, I., Iriyadi, D., Ice, Y. W., Susanti, H., & Nasution, R. A. (2022). A Rasch analysis of perceived stigma of covid-19 among nurses in Indonesia questionnaire. *Psychological Thought*, 15(1), 12.
- Falani, I., & Irdiyansyah, I. (2018). The ability of estimation stability and item parameter characteristics reviewed by item response theory model. *Advances in Social Science, Education and Humanities Research*, 15. <https://doi.org/https://doi.org/10.2991/icems-17.2018.34>
- Falani, I., Nisraeni, N., & Irdiyansyah, I. (2017). The ability of estimation stability and item parameter characteristics reviewed by Item Response Theory model. *International Conference on Education in Muslim Society (ICEMS 2017)*, 175–178.
- Farman, F., Anjelina, S., Putri, Q. T., Mardiah, N. A., & Sari, K. (2021). Pengembangan instrumen penilaian pembelajaran matematika berbasis ispring suite. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 10(4), 2040. <https://doi.org/10.24127/ajpm.v10i4.4288>
- Fatkhudin, A., Surarso, B., & Subagio, A. (2016). Item response theory model empat parameter logistik pada computerized adaptive test. *Jurnal Sistem Informasi Bisnis*, 4. <https://doi.org/10.21456/vol4iss2pp121-129>
- Fernanda, J. W., & Hidayah, N. (2020). Analisis kualitas soal ujian statistika menggunakan classical test theory dan rasch model. *Square: Journal of Mathematics and Mathematics Education*, 2(1), 49. <https://doi.org/10.21580/square.2020.2.1.5363>
- Fitri, D., Mujahidawati, M., & Sabil, H. (2025). Deskripsi media video pembelajaran menggunakan problem based learning untuk meningkatkan kemampuan berpikir kritis matematis. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, 5(2).
- Fitri, M., Mujahidawati, M., & Falani, I. (2025). Pengembangan e-modul etnomatematika 3d terintegrasi virtual reality berbasis tpack untuk meningkatkan kemampuan berpikir kritis siswa. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 14(2), 368. <https://doi.org/10.24127/ajpm.v14i2.10191>

- Frey, F. (2020). Test theory and classical test theory. *The International Encyclopedia of Media Psychology*, 1–6. <https://doi.org/10.1002/9781119011071.iemp0047>
- Friyatmi, F. (2018). Estimasi parameter tes dengan penskoran politomus menggunakan graded response model pada sampel kecil. *Jurnal Inovasi Pendidikan Ekonomi*, 8, 22. <https://doi.org/10.24036/01104490>
- Gustiawan, W., Niaga, A., & Padang, P. N. (2019). The theories of leadership: a review of papers. *Jurnal Ilmiah Poli Bisnis*, 11(1), 65–72.
- Habibah, S., & Marlina, R. (2023). Analisis kemampuan berpikir kritis matematis siswa kelas VIII SMP. *PHI: Jurnal Pendidikan Matematika*, 7(2), 141. <https://doi.org/10.33087/phi.v7i2.287>
- Halpern, D., & Dunn, D. (2022). *Thought and Knowledge: An Introduction to Critical Thinking*. New York, NY: Routledge. <https://doi.org/10.4324/9781003025412>
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory* (Issue v. 2). Sage Publications.
- Handayani, R., Destania, Y., & Muhammadiyah Bengkulu, U. (2021). Soal essay materi aritmatika sosial untuk kemampuan berpikir kritis matematis siswa kelas VII. *Indiktika : Jurnal Inovasi Pendidikan Matematika*, 4(1). <https://doi.org/https://doi.org/10.31851/indiktika.v4i1.6791>
- Hartanti, A., & Purnomo, Y. W. (2023). Peningkatan keterampilan berpikir kritis matematis dan hasil belajar menggunakan model pembelajaran problem based learning. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 12(2), 1728. <https://doi.org/10.24127/ajpm.v12i2.6891>
- Hayat, B. (2021). Klasika: program analisis item dan tes dengan pendekatan klasik. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia*, 10(1), 1–11. <https://doi.org/10.15408/jp3i.v10i1.20551>
- Huebner, A., & Skar, G. B. (2021). Conditional standard error of measurement: classical test theory, generalizability theory and many-facet rasch measurement with applications to writing assessment. *Practical Assessment, Research and Evaluation*, 26, 1–20. <https://doi.org/10.7275/vzmm-0z68>
- Indrawatiningsih, N., Purwanto, As'Ari, A. R., Dwiyan, Sudirman, & Rahardi, R. (2019). The ability of high school students' critical thinking in solving trigonometric problems. *IOP Conference Series: Earth and Environmental Science*, 243(1). <https://doi.org/10.1088/1755-1315/243/1/012050>

- Inganah, S., Darmayanti, R., & Rizki, N. (2023). Problems, solutions, and expectations: 6c integration of 21 st century education into learning mathematics. *JEMS (Journal of Mathematics and Science Education)*, 11(1), 220–238.
- Keliat, B. A., Nasution, R. A., Handini, I. T., & Falani, I. (2023). Psychometric properties evaluation of the Indonesian version of the Beck Scale for Suicide Ideation (BSSI) questionnaire using a Rasch model. *European Review for Medical & Pharmacological Sciences*, 27(21).
- Kogar, E. Y., & Kogar, H. (2018). Examination of dimensionality and latent trait scores on mixed-format tests. *PUPIL: International Journal of Teaching, Education and Learning*, 2(1), 29–49. <https://doi.org/10.20319/pijtel.2018.21.2949>
- Komala, P., Kaniawati, I., & Efendi, R. (2020). Karakterisasi soal tes keterampilan berpikir kritis menggunakan analisis item response theory pada materi fluida statis. *WaPFi (Wahana Pendidikan Fisika)*, 5(2), 100–109. <https://doi.org/10.17509/wapfi.v5i2.7755>
- Lee, W.-C., Kim, S. Y., Choi, J., & Kang, Y. (2020). IRT approaches to modeling scores on mixed-format tests. *Journal of Educational Measurement*, 57(2), 230–254. <https://doi.org/https://doi.org/10.1111/jedm.12248>
- Leniati, B., & Indarini, E. (2021). Keefektifan model pembelajaran kooperatif tipe jigsaw dan tsts (two stay two stray) terhadap kemampuan berpikir kritis pada pembelajaran matematika siswa. *Mimbar Ilmu*. <https://ejournal.undiksha.ac.id/index.php/MI/article/view/33359>
- Luo, Y. (2018). Parameter recovery with marginal maximum likelihood and markov chain monte carlo estimation for the generalized partial credit model. *National Center for Assessment in Saudi Arabia*, 1–36. <https://doi.org/https://doi.org/10.48550/arXiv.1809.07359>
- LW, A., DR, K., PW, A., KA, C., Mayer, R., PR, P., Rath, J., & MC, W. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*.
- Marsigit, & Susilo, N. B. (2006). *Matematika 1 SMP Kelas VII*. Cimahi: Quadra.
- Maulana, A. (2022). Analisis validitas, reliabilitas, dan kelayakan instrumen penilaian rasa percaya diri siswa. *Jurnal Kualita Pendidikan*, 3(3), 2774–2156. <https://doi.org/https://doi.org/10.51651/jkp.v3i3.331>
- Mertasari, N. M. S., & Candiasa, I. M. (2023). Development and validation of metacognitive test in programming using graded response model. *IConVET 2022: Proceedings of the 5th International Conference on Vocational Education and Technology, IConVET 2022, 6 October 2022, Singaraja, Bali, Indonesia*, 70.

- Mujahidawati, Novferma, Simatupang, G. M., Frianto, A., & Romundza, F. (2023). Analysis of university students' higher order thinking skills (HOTS) in the differential equation courses in the time of Covid-19 pandemic. *AIP Conference Proceedings*, 2698(1), 060047.
- Muraki, E. (1992). A generalized partial credit model: application of an em algorithm index terms: item response model, national assessment of educational progress, nominal response model, partial credit model, polytomous response model, rating scale model. *Applied Psychological Measurement*, 16(1), 159–176. <https://doi.org/https://hdl.handle.net/11299/115645>
- Mutmainna, M., Hasyim, M., Syamsuriwal, S., Setiaji, B., & Setiawati, F. (2024). Modeling of colorado learning attitude science survey in indonesian version: a study with applying item response theory. *JP3I (Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia)*, 13, 80–104. <https://doi.org/10.15408/jp3i.v13i1.36745>
- Naga, D. S. (2012). *Teori Sekor pada Pengukuran Mental*. Jakarta: PT Nagarani Citrayasa.
- Nasihudin, N., & Hariyadin, H. (2021). Pengembangan keterampilan dalam pembelajaran. *Jurnal Pendidikan Indonesia*, 2(04), 733–743. <https://doi.org/10.59141/japendi.v2i04.150>
- Nasution, R. A., & Falani, I. (2024). Measuring cyberbullying levels among adolescents in integrated islamic junior high schools in jambi city: an implementation of the rasch model. *Journal of Educational, Health & Community Psychology (JEHCP)*, 13(1).
- Niken, Somayasa, W., Budiman, H., Ruslan, & Adi Wibawa, G. N. (2024). Uji goodness of fit dengan statistik pulkstenis - robinson dalam model regresi logistik ordinal. *Jurnal Matematika Komputasi Dan Statistika*, 4(1), 571–576. <https://doi.org/10.33772/jmks.v4i1.72>
- Nufus, H., & Kusaeri, A. (2020). Analisis tingkat kemampuan berpikir kritis siswa dalam memecahkan masalah geometri. *Jurnal Pendidikan Matematika Indonesia*, 5(2), 49–55.
- Nuryadi, N., Sukestiyarno, Y., Suyitno, H., & Kharisudin, I. (2024). Mathematical critical thinking profile of students on pisa framework space and shape content questions reviewed from self-efficacy. *Educational Administration: Theory and Practice*, 30(4), 1380–1400. <https://doi.org/10.53555/kuvey.v30i4.1228>
- Osterlind, S. J. (2002). *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats* (2nd ed). New York: Kluwer Academic.

- Partono, P., Wardhani, H. N., Setyowati, N. I., Tsalitsa, A., & Putri, S. N. (2021). Strategi meningkatkan kompetensi 4c (critical thinking, creativity, communication, & collaborative). *Jurnal Penelitian Ilmu Pendidikan*, 14(1), 41–52. <https://doi.org/10.21831/jpipfip.v14i1.35810>
- Ramalisa, Y., Falani, I., & Pasaribu, F. T. (2023). Rasch analysis in developing Jambi culture-based ethnomathematics test for prospective mathematics teachers. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, 243–257.
- Retnawati, H. (2015). Perbandingan estimasi kemampuan laten antara metode maksimum likelihood dan metode bayes [latent capability between maximum likelihood method and bayes method]. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 19(2), 145–155.
- Rosidah, N. A., Ramalis, T. R., & Suyana, I. (2018). Karakteristik tes keterampilan berpikir kritis (KBK). *Jurnal Inovasi Dan Pembelajaran Fisika*, January 2018, 54–63.
- Santoso, A., Pardede, T., Djidu, H., Apino, E., Rafi, I., Rosyada, M. N., & Abd Hamid, H. S. (2022). The effect of scoring correction and model fit on the estimation of ability parameter and person fit on polytomous item response theory. *REID (Research and Evaluation in Education)*, 8(2), 140–151. <https://doi.org/10.21831/reid.v8i2.54429>
- Sappaile, B. I. (2010). Konsep penelitian ex-post facto. *Jurnal Pendidikan Matematika*, 1(2), 105–113. <https://doi.org/10.36709/jpm.v1i2.1946>
- Saputri, H. A., Zulhijrah, Larasati, N. J., & Shaleh. (2023). Analisis instrumen assesmen: validitas, reliabilitas, tingkat kesukaran, dan daya beda butir soal. *Didaktik : Jurnal Ilmiah PGSD FKIP Universitas Mandiri*, 09(05), 2986–2995.
- Satria, M. R. (2024). Transformasi standar penilaian pendidikan dan revitalisasi penilaian pembelajaran di indonesia. *Jurnal Penelitian Kebijakan Pendidikan*, 17(1), 57–66. <https://doi.org/10.24832/jpkp.v17i1.930>
- Septiani, A. N., Pratiwi, D., & Rossy, R. (2023). Evaluasi pembelajaran dalam peningkatan mutu pendidikan di sekolah dasar. *Masaliq*, 3(5), 824–832. <https://doi.org/10.58578/masaliq.v3i5.1380>
- Setiana, D. S., & Purwoko, R. Y. (2020). Analisis kemampuan berpikir kritis ditinjau dari gaya belajar matematika siswa. *Jurnal Riset Pendidikan Matematika*, 7(2), 163–177. <https://doi.org/10.21831/jrpm.v7i2.34290>

- Siburian, J., Sinaga, E., & Murni, P. (2023). Kemampuan berpikir kritis melalui implementasi flipped classroom pada siswa sma. *INKUIRI: Jurnal Pendidikan IPA*, 12(1), 71. <https://doi.org/10.20961/inkuiri.v12i1.68213>
- Sugianto, R., Cholily, Y. M., Darmayanti, R., Rahmah, K., & Hasanah, N. (2022). Development of rainbow mathematics card in tgt learning for increasing mathematics communication ability. *Kreano, Jurnal Matematika Kreatif-Inovatif*, 13(2), 221–233. <https://doi.org/10.15294/kreano.v13i2.38068>
- Sugiyono. (2015). *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)*. Bandung: Alfabeta.
- Sugiyono. (2019). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*. Jakarta: Trim Komunikata Publishing House.
- Supriyati, Y., Falani, I., & Maulana, S. (2021). The information function of mixed-format test of physics learning outcomes measurement. *AIP Conference Proceedings*, 2320(1).
- Susanti, D., & Hernawati, C. (2022). Keterampilan berpikir kritis mahasiswa dalam menyelesaikan soal higher order thinking skills. *EDU-MAT: Jurnal Pendidikan Matematika*, 10(1), 115. <https://doi.org/10.20527/edumat.v10i1.12190>
- Tao, Jian, Shi, Ning-Zhong, & Chang, Hua-Hua. (2012). Item-weighted likelihood method for ability estimation in tests composed of both dichotomous and polytomous items. *Journal of Educational and Behavioral Statistics*, 37(2), 298–315. <https://doi.org/10.3102/1076998610393969>
- Teluma, M., & Rivaie, W. (2019). *Penilaian* (1st ed.). PGRI Prov Kalbar.
- Wang, S., Asi, C., & Wang, T. (2002). *Relative precision of ability estimation in polytomous CAT: A comparison under the generalized partial credit model and graded response model*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. <https://eric.ed.gov/?id=ED477926>
- Wisnu Deni Arif, Tri Rijanto, Rina Harimurti, & Yulia Fransisca. (2023). Penerapan Partial Credit Model (PCM) untuk meningkatkan kemampuan berpikir kritis siswa pada mata pelajaran perbaikan peralatan listrik pada kelas XI di SMK Negeri 1 Kediri. *Jurnal Elektronika Dan Teknik Informatika Terapan (JENTIK)*, 1(3), 249–262. <https://doi.org/10.59061/jentik.v1i3.405>

- Yilmaz, H. B. (2019). A comparison of irt model combinations for assessing fit in a mixed format elementary school science test. *International Electronic Journal of Elementary Education*, 11(5 SE-Articles), 539–545.
- Yokia Marsita, & Ishaq Nuriadin. (2024). Analisis kemampuan berpikir kritis matematis siswa dalam menyelesaikan soal cerita spldv ditinjau dari disposisi matematis. *Jurnal Math-UMB.EDU*, 11(2), 129–136. <https://doi.org/10.36085/mathumbedu.v11i2.6143>
- Zahri, T. N., Yusuf, A. M., & S, N. (2017). Hubungan gaya belajar dan keterampilan belajar dengan hasil belajar mahasiswa serta implikasinya dalam pelayanan Bimbingan dan Konseling di Fakultas Ilmu Pendidikan Universitas Negeri Padang. *Konselor*, 6(1), 18. <https://doi.org/10.24036/02017615734-0-00>
- Zain, A. Z., Ramalis, T. R., & Muslim, M. (2022). Karakterisasi instrumen tes keterampilan berpikir kreatif berdasarkan analisis partial credit model. *Jurnal Ilmiah Pendidikan Fisika*, 6(1), 176. <https://doi.org/10.20527/jipf.v6i1.4806>
- Zainal, N. F. (2020). Pengukuran, assessment dan evaluasi dalam pembelajaran matematika. *Laplace: Jurnal Pendidikan Matematika*, 3(1), 8–26. <https://doi.org/10.31537/laplace.v3i1.310>
- Zohar, A., & Dori, Y. J. (2003). Higher order thinking skills and low-achieving students: are they mutually exclusive?. *Journal of the Learning Sciences*, 12(2), 145–181. https://doi.org/10.1207/S15327809JLS1202_1