

DAFTAR PUSTAKA

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- 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.
- Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. Phi Delta Kappan, 80(2), 139–148.
- Bond, T. G., Fox, C. M., & Yan, Z. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences (4th ed.)*. Routledge.
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer.
- Chimayah, M., Windayani, N., Kurniasih, N., Subarkah, C. Z., & Sukmawardani, Y. (2025). Analysis of students misconceptions using instruments three-tier diagnostic test on acid-base material. *JISIP: Jurnal Ilmu Sosial dan Pendidikan*, 9(1), 408–415. <https://doi.org/10.58258/jisip.v9i1.8215>
- Couch, B. A., Hubbard, J. K., & Brassil, C. E. (2018). Multiple–true–false questions reveal the limits of the multiple–choice format for detecting students with incomplete understandings. *BioScience*, 68(6), 455–463. <https://doi.org/10.1093/biosci/biy037>
- Djarwo, C. F., & Kafiari, F. P. (2023). Analyzing misconceptions of acid-base topic among chemistry education students in online learning settings: A case study. *Hydrogen: Jurnal Kependidikan Kimia*, 11(2), 202–208. <https://doi.org/10.33394/hjkk.v11i2.7147>
- Fisher, W. P., Jr. (1992). Reliability statistics. *Rasch Measurement Transactions*, 6(3), 238.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction (8th ed.)*. Pearson/Allyn & Bacon.
- Haladyna, T. M. (1993). *Developing and validating multiple-choice test items*. Lawrence Erlbaum Associates.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items (3rd ed.)*. Lawrence Erlbaum Associates.
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge.
- Hattie, J., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning*, 1, 16013. <https://doi.org/10.1038/npjscilearn.2016.13>
- Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Lawrence Erlbaum Associates.

- Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, 7(2), 75–83. <https://doi.org/10.1111/j.1365-2729.1991.tb00230.x>
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. <https://doi.org/10.1111/jedm.12000>
- Kementerian Pendidikan Dasar dan Menengah Republik Indonesia. (2025). Peraturan Menteri Pendidikan Dasar dan Menengah Nomor 13 Tahun 2025 tentang Perubahan atas Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 12 Tahun 2024 tentang Kurikulum pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, dan Jenjang Pendidikan Menengah.
- Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan. (2022). Keputusan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Nomor 033/H/KR/2022 tentang Perubahan atas Keputusan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Nomor 008/H/KR/2022 tentang Capaian Pembelajaran pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, dan Jenjang Pendidikan Menengah pada Kurikulum Merdeka. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.
- Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan. (2025). Keputusan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Nomor 046/H/KR/2025 tentang Capaian Pembelajaran pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, dan Jenjang Pendidikan Menengah. Kementerian Pendidikan Dasar dan Menengah.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, 7(4), 328.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149–174. <https://doi.org/10.1007/BF02296272>
- Meyer, J. P. (2014). *Applied measurement with jMetrik*. Routledge.
- Mubarokah, F. D., Mulyani, S., & Indriyanti, N. Y. (2018). Identifying students' misconceptions of acid-base concepts using a three-tier diagnostic test: A case of Indonesia and Thailand. *Journal of Turkish Science Education*, 15(Special Issue), 51–58.
- Nakhleh, M. B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education*, 69(3), 191–196. <https://doi.org/10.1021/ed069p191>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.)*. McGraw-Hill.
- OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>

- Pan, H., & Henriques, L. (2015). Students' alternate conceptions on acids and bases. *School Science and Mathematics*, 115(5), 237–243. <https://doi.org/10.1111/ssm.12124>
- Pınarbaşı, T. (2007). Turkish undergraduate students' misconceptions on acids and bases. *Journal of Baltic Science Education*, 6(1), 23–34.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Putriawati, W., & Chiani, S. H. (2025). Development of an instrument to map acid–base misconceptions among high-achieving senior high school students in Eastern Indonesia. *Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram*, 13(4), 1186–1198. <https://doi.org/10.33394/j-ps.v13i4.17646>
- Rakhmalinda, F., Arifiani, R., Puspawati, W., & Simamora, A. M. (2024). Analysis of students' misconceptions on acid-base topic using the two-tier diagnostic test. *Journal of Computers for Science and Mathematics Learning*, 1(1), 27–43. <https://doi.org/10.70232/9xp1nt17>
- Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. *Danish Institute for Educational Research*.
- Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346–362. <https://doi.org/10.1037/0022-0663.93.2.346>
- Rodriguez, M. C. (1997). The art & science of item writing: A meta-analysis of multiple-choice item format effects. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Romine, W. L., Todd, A. N., & Clark, T. B. (2016). How do undergraduate students conceptualize acid–base chemistry? Measurement of a concept progression. *Science Education*, 100(6), 1030–1050. <https://doi.org/10.1002/sce.21240>
- Ross, B., & Munby, H. (1991). Concept mapping and misconceptions: A study of high-school students' understandings of acids and bases. *International Journal of Science Education*, 13(1), 11–23. <https://doi.org/10.1080/0950069910130102>
- Shaafi, N. F., Yusof, M. M. M., Ellianawati, E., Subali, B., & Raji'e, M. H. H. (2025). Investigating misconceptions about acids and bases among pre-service science teachers. *Journal of Education and Learning (EduLearn)*, 19(1), 460–477.
- Taber, K. S. (2002). Chemical misconceptions – Prevention, diagnosis and cure: Theoretical background (Vol. 1). *Royal Society of Chemistry*.
- Taber, K. S. (2019). Alternative conceptions and the learning of chemistry. *Israel Journal of Chemistry*, 59(6–7), 450–469. <https://doi.org/10.1002/ijch.201800046>

- Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). Instructional development for training teachers of exceptional children: A sourcebook. Indiana University.
- Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education*, 91(9), 1426–1431. <https://doi.org/10.1021/ed500076x>
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159–169. <https://doi.org/10.1080/0950069880100204>
- Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2003). The role of submicroscopic and symbolic representations in chemical explanations. *International Journal of Science Education*, 25(11), 1353–1368. <https://doi.org/10.1080/0950069032000070306>

