

## Daftar Pustaka

- Abedalaziz, N. (2010). Detecting gender related DIF using Logistic Regression and Mantel-Haenszel approaches. *Procedia - Social and Behavioral Sciences*, 7, 406–413. <https://doi.org/10.1016/j.sbspro.2010.10.055>
- Alavi, S., Ali Rezaee, A., Amirian, S. (2011). Academic discipline DIF in an english language proficiency test. *Journal of English Language Teaching and Learning*, 3(7), 39-65.
- American Educational Research Association (Ed.). (2014). *Standards for educational and psychological testing*. Washington, D.C: American Educational Research Association.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7 edition). Upper Saddle River, N.J: Pearson.
- Angoff, W. H. (1993). Perspectives on differential item functioning methodology. Dalam *Differential item functioning* (hlm. 3–23). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Arnani, M. (2019). Passing grade CPNS 2019 lebih rendah, berapa rincian nilainya?. Diunduh dari <https://www.kompas.com/tren/read/2019/11/14/153511865/passing-grade-cpns-2019-lebih-rendah-berapa-rincian-nilainya?page=all> tanggal 2 November 2019
- Azwar, S. (2016a). *Konstruksi tes kemampuan kognitif*. Pustaka Pelajar: Yogyakarta.
- Azwar, S. (2016b). *Dasar-dasar psikometrika*. Pustaka Pelajar: Yogyakarta.
- Bartholomew, D. (2004). Measuring intelligence: Facts and fallacies. *Measuring Intelligence: Facts and Fallacies*, 1–172. <https://doi.org/10.1017/CBO9780511490019>
- Bencsik, A., Juhász, T., & Horváth-Csikós, G. (2016). Y and Z Generations at Workplaces. *Journal of Competitiveness*, 6, 90–106. <https://doi.org/10.7441/joc.2016.03.06>
- Berberoglu, G., & Tosunoglu, C. (1995). Exploratory and confirmatory factor analyses of an Environmental Attitude Scale (EAS) for Turkish University students. *The Journal of Environmental Education*, 26(3), 40–43. <https://doi.org/10.1080/00958964.1995.9941444>
- Bicer, A., & Navruz, B. (2015). Stem schools Vs. Non-stem schools: Comparing students' mathematics growth rate on high-stakes test performance. *International Journal on New Trends in Education and Their Implications*. 6(1), 13, 138-150.

- Brown, R. T., Reynolds, C. R., Whitaker, J. S. (1999). Bias in mental testing since bias in mental testing. *School Psychology Quarterly*, Vol. 14, No. 3, 1999, pp. 208-238
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks [Calif.]: Sage Publications.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge Univ. Press.
- Costley, K. C. (2014). *The positive effects of technology on teaching and student learning*. Associate Professor of Curriculum & Instruction Arkansas Tech University, Arkansas.
- Craig, B. (2017). *The empirical selection of anchor items using a multistage approach*. Graduate Theses and Dissertations. <http://scholarcommons.usf.edu/etd/6819>
- Cunningham, T. J., Berkman, L. F., Gortmaker, S. L., Kiefe, C. I., Jacobs, D. R., Seeman, T. E., & Kawachi, I. (2011). Assessment of differential item functioning in the experiences of discrimination index. *American Journal of Epidemiology*, 174(11), 1266–1274. <https://doi.org/10.1093/aje/kwr253>
- Desjardins, C. D., & Bulut, O. (2018). *Handbook of educational measurement and psychometrics using R* (1 ed.). CRC Press: London. <https://doi.org/10.1201/b20498>
- Elena Oliveri, M., Lawless, R., Robin, F., & Bridgeman, B. (2018). An exploratory analysis of differential item functioning and its possible sources in a higher education admissions context. *Applied Measurement in Education*, 31(1), 1–16. <https://doi.org/10.1080/08957347.2017.1391258>
- Erdem, A. (2015). A simultaneous approach to inverse source problem by Green's function. *Mathematical Methods in the Applied Sciences*, 38(7), 1393–1404. <https://doi.org/10.1002/mma.3154>
- Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000). Sex differences in early verbal and non-verbal cognitive development. *Developmental Science*, 3(2), 206–215. <https://doi.org/10.1111/1467-7687.00114>
- Gómez-Benito, J., Sireci, S., Padilla, J., Hidalgo, M. D., & Benítez, I. (2018). Differential item functioning: beyond validity evidence based on internal structure. *Psicothema* 2018, Vol. 30, No. 1, 104-109. doi: 10.7334/psicothema2017.183
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24(1), 79–132. [https://doi.org/10.1016/S0160-2896\(97\)90014-3](https://doi.org/10.1016/S0160-2896(97)90014-3)

- Gregory, M. R. (2002). Constructivism, standards, and the classroom community of inquiry. *Educational Theory*, 52(4), 397–408. <https://doi.org/10.1111/j.1741-5446.2002.00397.x>
- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269(5220), 41–45. <https://doi.org/10.1126/science.7604277>
- Holland, P. W., & Thayer, D. T. (1986). Differential item functioning and the Mantel-Haenszel procedure. *ETS Research Report Series*, 1986(2), i–24. <https://doi.org/10.1002/j.2330-8516.1986.tb00186.x>
- Jensen, . R. (1980). *Bias in mental testing*. New York: Free Press.
- Jurkiewicz, C. L. (2000). Generation x and the public employee. *Public Personnel Management*, 29(1), 55–74. <https://doi.org/10.1177/009102600002900105>
- Kagan, J. (1971). *Change and continuity in infancy*. Oxford, England: John Wiley & Sons.
- Katz-Wise, S. L., Priess, H. A., & Hyde, J. S. (2010). Gender-role attitudes and behavior across the transition to parenthood. *Developmental Psychology*, 46(1), 18–28. <https://doi.org/10.1037/a0017820>
- Kemendikbud dan Casio Kerja Sama Tingkatkan Kompetensi Pendidik dan Kurikulum. (2018, September). Kementrian Pendidikan dan Kebudayaan. Diunduh dari <https://www.kemdikbud.go.id/main/blog/2018/09/kemendikbud-dan-casio-kerja-sama-tingkatkan-kompetensi-pendidik-dan-kurikulum> tanggal 5 Agustus 2019.
- Kline, P. (1999). *Handbook of psychological testing, second edition* (2 edition). London ; New York: Routledge.
- Kuncel, N. R., & Hezlett, S. A. (2007). Assessment: standardized tests predict graduate students' success. *Science*, 315(5815), 1080–1081. <https://doi.org/10.1126/science.1136618>
- Kuncel, Nathan R., & Hezlett, S. A. (2010). Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current Directions in Psychological Science*, 19(6), 339–345. <https://doi.org/10.1177/0963721410389459>
- Lawrence, I. M., Curley, W. E., & McHale, F. J. (1988). Differential item functioning for males and females on sat®-verbal reading subscore items. *ETS Research Report Series*, 1988(1), i–55. <https://doi.org/10.1002/j.2330-8516.1988.tb00266.x>
- Linden, W. J. van der, & Hambleton, R. K. (1997). *Handbook of modern Item Response Theory*. Springer: New York

- Liu, O. L., dan Wilson, M. (2009). Gender differences in large-scale math assessments: PISA Trend 2000 and 2003. *Applied Measurement in Education*, 22(2), 164–184. <https://doi.org/10.1080/08957340902754635>
- Lohman, D. F., dan Lakin, J. M. (2009). Reasoning and intelligence. Dalam Stenberg, R. J., dan Kaufman, S. B. *Handbook of Intelligence* (2nd ed.). New York: Cambridge University Press.
- Lopez, G. E. (2012). *Detection and classification of dif types using parametric and nonparametric methods: A comparison of the IRT-Likelihood Ratio Test, Crossing-SIBTEST, and Logistic Regression Procedures*. Graduate Theses and Dissertations. <http://scholarcommons.usf.edu/etd/4131>
- Lynn, R., & Mikk, J. (2009). Sex differences in reading achievement. *Trames. Journal of the Humanities and Social Sciences*, 13(1), 3. <https://doi.org/10.3176/tr.2009.1.01>
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *JNCI: Journal of the National Cancer Institute*, 22(4), 719–748. <https://doi.org/10.1093/jnci/22.4.719>
- Martinková, P., Drabinová, A., Liaw, Y.-L., Sanders, E. A., McFarland, J. L., & Price, R. M. (2017). Checking equity: why differential item functioning analysis should be a routine part of developing conceptual assessments. *CBE—Life Sciences Education*, 16(2), rm2. <https://doi.org/10.1187/cbe.16-10-0307>
- McPeck, W. M., & Wild, C. L. (1992). Identifying differentially functioning items in the nte core battery. *ETS Research Report Series*, 1992(2), i–44. <https://doi.org/10.1002/j.2333-8504.1992.tb01493.x>
- Mellenbergh, G. J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics*, 7(2), 105–118. <https://doi.org/10.3102/10769986007002105>
- O'Neill, K. A., McPeck, W. M., & Wild, C. L. (1993). Differential item functioning on the graduate management admission test. *ETS Research Report Series*, 1993(2), i–46. <https://doi.org/10.1002/j.2333-8504.1993.tb01546.x>
- Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Cognitive ability in selection decisions. *Handbook of Understanding and Measuring Intelligence*, 431–468. <https://doi.org/10.4135/9781452233529.n24>
- Osterlind, S. J. (1998). *Constructing test items: Multiple-choice, constructed-response, performance, and other formats* (2nd ed). Kluwer Academic Publishers: New York.
- Osterlind, S. J., & Everson, H. T. (2009). *Differential item functioning*. SAGE Publications: New York.

- Pae, T.-I. (2004). DIF for examinees with different academic backgrounds. *Language Testing*, 21(1), 53–73.  
<https://doi.org/10.1191/0265532204lt274oa>
- Piraksa, C., Srisawasdi, N., & Koul, R. (2014). Effect of gender on student's scientific reasoning ability: a case study in Thailand. *Procedia - Social and Behavioral Sciences*, 116, 486–491.  
<https://doi.org/10.1016/j.sbspro.2014.01.245>
- Putra, Y. S. (2016). Theoretical review : teori perbedaan generasi. *Jurnal Ilmiah Among Makarti*. Vol.9, No. 18, 123-134.
- Ramos, H., Ramos, S. J., & Juntune, J. E. (2018). Test review: test of General Reasoning Ability (TOGRA), by C. R. Reynolds. *Journal of Psychoeducational Assessment*, 36(8), 869–872.  
<https://doi.org/10.1177/0734282917709741>
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York, NY: Routledge.
- Reise, S. P., & Revicki, D. A. (Ed.). (2014). *Handbook of item response theory modeling: applications to typical performance assessment* (1 edition). New York: Routledge.
- Reynolds, M. R., Scheiber, C., Hajovsky, D. B., Schwartz, B., & Kaufman, A. S. (2015). Gender differences in academic achievement: is writing an exception to the gender similarities hypothesis? *The Journal of Genetic Psychology*, 176(4), 211–234.  
<https://doi.org/10.1080/00221325.2015.1036833>
- Rogers, H. J., & Swaminathan, H. (1993). A comparison of logistic regression and mantel-haenszel procedures for detecting differential item functioning. *Applied Psychological Measurement*, 17(2), 105–116.  
<https://doi.org/10.1177/014662169301700201>
- Runnels, J. (2013). Measuring differential item and test functioning across academic disciplines. *Language Testing in Asia*, 3(1), 9.  
<https://doi.org/10.1186/2229-0443-3-9>
- Santoso, A. (2018). *Equivalence testing for anchor selection in differential item functioning (DIF) detection* (University Of Notre Dame). Diambil dari <https://curate.nd.edu/show/2r36tx34933>
- Scheuneman, J. D., & Gerritz, K. (1990). Using differential item functioning procedures to explore sources of item difficulty and group performance characteristics. *Journal of Educational Measurement*, 27(2), 109–131.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274.  
<https://doi.org/10.1037/0033-2909.124.2.262>

- Schmitt, A. P., & Bleistein, C. A. (1987). Factors affecting differential item functioning for black examinees on scholastic aptitude test analogy items. *ETS Research Report Series*, 1987(1), i–46. <https://doi.org/10.1002/j.2330-8516.1987.tb00227.x>
- Shepard, L., Camilli, G. & Averil, M. (1981). Comparison of procedures for detecting test-item bias with both internal and external ability criteria. *Journal of Educational Statistics*, 6, 317-375.
- Stark, S., Chernyshenko, O. S., Chan, K.-Y., Lee, W. C., & Drasgow, F. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of Applied Psychology*, 86(5), 943–953. <https://doi.org/10.1037/0021-9010.86.5.943>
- Steinmayr, R., Bergold, S., Margraf-Stiksrud, J., & Freund, P. A. (2015). Gender differences on general knowledge tests: Are they due to Differential Item Functioning? *Intelligence*, 50, 164–174. <https://doi.org/10.1016/j.intell.2015.04.001>
- Sudaryono. (2012). Sensitivity of differential item functioning (DIF) detection method. *Jurnal Evaluasi Pendidikan*, Vol. 3, No. 1, 82-94
- Suryabrata, S. (1998). *Pengembangan alat ukur psikologi*. Yogyakarta: Andi
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361–370.
- Takeuchi, H., Taki, Y., Sekiguchi, A., Nouchi, R., Kotozaki, Y., Nakagawa, S., ... Kawashima, R. (2015). Brain structures in the sciences and humanities. *Brain Structure & Function*, 220(6), 3295–3305. <https://doi.org/10.1007/s00429-014-0857-y>
- Tan, R., & Dejas, A. W. (2019). Comparing problem solving ability of stem and non-stem entrants to bachelor of science in mathematics education program. *Sci.Int.(Lahore)*, 31(1)B, 5-7.
- Thissen, D., Steinberg, L., & Wainer, H. (1988). Use of item response theory in the study of group differences in trace lines. Dalam *Test validity* (hlm. 147–172). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. Dalam *Differential item functioning* (hlm. 67–113). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Urbina, S. (2014). *Essentials of psychological testing (2 edition)*. Hoboken, New Jersey: Wiley.



- Wedman, J. (2018). Reasons for gender-related differential item functioning in a college admissions test. *Scandinavian Journal of Educational Research*, 62(6), 959–970. <https://doi.org/10.1080/00313831.2017.1402365>
- Zoref, L., & Williams, P. (1980). A look at content bias in iq tests. *Journal of Educational Measurement*, 17(4), 313–322.