Daftar Pustaka

Alfonsson, S., Maathz, P., & Hursti, T. (2014). Interformat reliability of digital psychiatric self-report questionnaires: a systematic review. *Journal of medical Internet research*, *16*(12), e268.

Asunta, P., Viholainen, H., Ahonen, T. et al. (2019). Psychometric properties of observational tools for identifying motor difficulties – a systematic review. *BMC Pediatr* 19, 322. https://doi.org/10.1186/s12887-019-1657-6

Azwar, S. (2008). Kualitas Tes Potensi Akademik Versi 07A. *Jurnal Penelitian dan Evaluasi Pendidikan, 12(2), 232-250*.

Azwar, S. (2019). *Konstruksi Tes Kemampuan Kognitif*. Yogyakarta: Pustaka Pelajar

Baker, F. B., & Kim, S. H. (2017). *The basics of item response theory using R*. New York: Springer.

Barak, A. (2011). Internet-based psychological testing and assessment. In *Daring counseling* (pp. 225-255). Academic Press.

Bock, R. D., & Gibbons, R. D. (2021). *Item response theory*. John Wiley & Sons.

Buchanan, T. (2002). Daring assessment: Desirable or dangerous?. *Professional psychology: Research and practice*, 33(2), 148. http://dx.doi.org/10.1037/0735-7028.33.2.148.

Buckley, J., Seery, N., & Canty, D. (2016). The validity and reliability of daring testing for the assessment of spatial ability. In *ASEE Engineering Design Graphics Division 71st Mid-Year Conference*. ASEE.

Bunderson, C. Victor; Inouye, Dillon K.; Olsen, James B. (1988). THE FOUR GENERATIONS OF COMPUTERIZED EDUCATIONAL MEASUREMENT. *ETS Research Report Series*, 1988(1), i–148. doi:10.1002/j.2330-8516.1988.tb00291.x

Chen, Y., Li, X., Liu, J., & Ying, Z. (2021). Item Response Theory--A Statistical Framework for Educational and Psychological Measurement. *arXiv preprint arXiv:2108.08604*.

Cooper, C. (2018). *Psychological testing: Theory and practice*. New York: Routledge.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98–104.

Costa, P., & Ferrão, M. E. (2015). On the complementarity of classical test theory and item response models: item difficulty estimates and computerized adaptive testing. *Ensaio: Avaliação e Políticas Públicas em Educação*, 23, 593-610.

Doherty, P. B. (2006). The context and culture of the Web research environment. *In Daring assessment, measurement and evaluation: Emerging practices* (pp. 10-27). IGI Global.

Furr R.M., and Bacharach, V.R. (2013). *Psychometrics. An Introduction (2nd ed.).* Los Angelers: Sage.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.).* Boston: Allyn & Bacon.

Ginty A.T. (2013) Psychometric Properties. In: Gellman M.D., Turner J.R. (eds) Encyclopedia of Behavioral Medicine. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-1005-9_480

Halim, D. (2020, March 24). Upaya Galakkan Social Distancing, Bekerja dari Rumah hingga Jeratan Pidana (Krisiandi, Ed.). Retrieved April 17, 2021, from https://nasional.kompas.com/read/2020/03/24/08215031/upaya-galakkan-social-distancing-bekerja-dari-rumah-hingga-jeratan-pidana

Hambleton, R.K. and Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Boston, MA: Kluwer-Nijhoff.

Hewson, C., Charlton, J., & Brosnan, M. (2007). Comparing daring and luring administration of multiple choice question assessments to psychology undergraduates: do assessment modality or computer attitudes influence performance?. *Psychology Learning & Teaching*, 6(1), 37-46.

Hu, Z., Lin, L., Wang, Y., & Li, J. (2021). The Integration of Classical Testing Theory and Item Response Theory. Psychology, 12(9), 1397-1409.

Jin, K. Y., Siu, W. L., & Huang, X. (2022). Exploring the Impact of Random Guessing in Distractor Analysis. *Journal of Educational Measurement.*

Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of research in personality*, 39(1), 103e129. http://dx.doi.org/10.1016/j.jrp.2004.09.009.

Kabacoff, R.I. (2015). *R in action: Data analysis and graphics with R (2nd Ed.).* Shelter Island, NY: Manning

Kaplan, R. M., & Saccuzzo, D. P. (2017). *Psychological Testing: Principles, Applications, and Issues* (Edisi 9). Belmont: Thomson Wadsworth.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior, *Proceedings of the National Academy of Sciences of the USA*, April 2013.

Kurniasari, E. (2020, November 17). UPAP Adakan Sosialisasi Daring Terkait Tes PAPs. Retrieved July 15, 2021, from https://psikologi.ugm.ac.id/upap-adakan-sosialisasi-daring-terkait-tes-paps/

LEE, J. W. (2020). Impact of Proctoring Environments on Student Performance: Daring vs Luring Proctored Exams. *The Journal of Asian Finance, Economics, and Business*, 7(8), 653-660.

Mair, P. (2018). *Modern psychometrics with R*. Cham, Switzerland: Springer

Marastuti, A., Anggoro, W. J., Marvianto, R. D., & Al Afghani, A. A. (2020). Perbandingan Properti Psikometri antara Tes PAPs Berbentuk Computer-Based dan Paper and Pencil Test. *Gadjah Mada Journal of Psychology (GamaJoP)*, *6*(1), 12-28.

McDonald, A. (2002). The Impact of Individual Differences on the Equivalence of ComputerBased and Paper-and-Pencil Educational Assessments. *Computers and Education*, 39(3), 299–312

Menteri Pendidikan. (2020). Surat Edaran Nomor 3 Tahun 2020 Tentang Pelaksanaan Pendidikan dalam Masa Darurat CoronaVirus (COVID-19).

Morgan, C., & O'Reilly, M. (2005). Ten Key Qualities of Asessment Daring. *In Daring assessment and measurement: Foundations and challenges* (pp. 86-101). IGI Global.

Matsunaga, M. (2008). Item parceling in Structural Equation Modeling: A primer. *Communication Methods and Measures,* 2(4), 260–293. doi: 10.1080/19312450802458935

Munoz, A., & Mackay, J. (2019). An daring testing design choice typology towards cheating threat minimisation. *Journal of University Teaching and Learning Practice*, 16(3), 5.

Naus, M. J., Philipp, L. M., & Samsi, M. (2009). From paper to pixels: A comparison of paper and computer formats in psychological assessment. *Computers in Human Behavior*, 25(1), 1e7. http://dx.doi.org/10.1016/j.chb.2008.05.012.

Noorbehbahani, F., Mohammadi, A., & Aminazadeh, M. (2022). A systematic review of research on cheating in daring exams from 2010 to 2021. *Education and Information Technologies*, 1-48.

Nunnally, J. C. (1978). *Psychometric theory*. New York : McGraw-Hill, c1978. 2d ed.

Nurcahyo, F. A. (2016). Aplikasi IRT dalam analisis aitem tes kognitif. *Buletin Psikologi*, 24(2), 64-75.

Obinne, A. D. E. (2012). Using IRT in Determining Test Item Prone to Guessing. *World Journal of Education*, 2(1), 91-95.

Olsen, J. B. (2006). Performance testing: validity issues and design considerations for daring testing. *In Daring assessment, measurement and evaluation: emerging practices* (pp. 259-274). IGI Global.

Paek, I., & Cole, K. (2020). *Using R for item response theory model applications*. New York: Routledge, Taylor & Francis Group.

Pritchard, A. E., Stephan, C. M., Zabel, T. A., & Jacobson, L. A. (2017). Is this the wave of the future? Examining the psychometric properties of child behavior ratings administered daring. *Computers in Human Behavior*, 70, 518-522.

Russell, A. E., Ford, T., & Russell, G. (2015). Socioeconomic associations with ADHD: Findings from a mediation analysis. *PLoS One*, 10(6). http://dx.doi.org/10.1371/journal.pone.0128248.

Rust, J., Kosiński, M., & Stillwell, D. (2021). *Modern psychometrics: The science of psychological assessment*. London: Routledge, Taylor & Francis Group.

Salkind, N. J. (2007). *Encyclopedia of measurement and statistics volume 1*. California: SAGE Publications, Inc.

Shum, D. H., O'Gorman, J., Myors, B., & Creed, P. (2017). *Psychological testing and assessment 3ʳᵈ ed*. Oxford University Press.

Stillwell, D. (2018, May). *MYPERSONALITY PROJECT*. myPersonality.org. Retrieved August 19, 2021, from https://sites.google.com/michalkosinski.com/mypersonality

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in science education*, 48(6), 1273-1296.

UPAP. (2019). *Buku Manual Tes Potensi Akademik Pascasarjana (PAPS)*. Yogyakarta: UPAP Fakultas Psikologi UGM

Wenzel, K., & Reinhard, M. A. (2020). Tests and academic cheating: do learning tasks influence cheating by way of negative evaluations?. *Social Psychology of Education*, 23(3), 721-753.

Wiberg, M., Lyrén, P. E., & Lind Pantzare, A. (2021). Schools, Universities and Large-Scale Assessment Responses to COVID-19: The Swedish Example. *Education Sciences*, 11(4), 175.

Widhiarso, W. & Suhapti, R. (2018). Penggunaan testlet dalam pengembangan tes psikologi. *Insan Jurnal Psikologi dan Kesehatan Mental*, 3(1).

Widhiarso, W. (2019). Pembuktian Validitas terkait Struktur Tes Potensi Akademik Pascasarjana (PAPS) Universitas Gadjah Mada. *Jurnal Psikologi*, 46(2), 145-162.

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior research methods*, 51(1), 409-428.

Ximénez, C., Maydeu-Olivares, A., Shi, D., & Revuelta, J. (2022). Assessing cutoff values of SEM fit indices: Advantages of the unbiased SRMR index and its cutoff criterion based on communality. *Structural Equation Modeling: A Multidisciplinary Journal*, 1-13.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans, *Proceedings of the National Academy of Sciences of the USA (PNAS),* September 2013.