

## DAFTAR PUSTAKA

- Afoakwa, E. O. (2004). *Application of response surface methodology for studying the product characteristics of extruded rice-cowpea-groundnut blends*. (August). <https://doi.org/10.1080/09637480400003238>
- Andrade, H. L., Du, Y., & Mycek, K. (n.d.). *Assessment in Education: Principles , Policy & Pctice Rubric - referenced self - assessment and middle school students ' writing*. (October 2014), 37–41. <https://doi.org/10.1080/09695941003696172>
- Andrich, D. (1978a). Application of a Psychometric Rating Model to Ordered Categories Which Are Scored with Successive Integers. *Applied Psychological Measurement*, 2(4), 581–594. <https://doi.org/10.1177/014662167800200413>
- Andrich, D. (1978b). Relationships Between the Thurstone and Rasch Approaches to Item Scaling. *Applied Psychological Measurement*, 2(3), 451–462. <https://doi.org/10.1177/014662167800200319>
- Andrich, D. (1978c). Scaling attitude items constructed and scored in the likert tradition. *Educational and Psychological Measurement*, 38(3), 665–680. <https://doi.org/10.1177/001316447803800308>
- Anneliese A. Singh; Lauren Lukkarila. (2017). *Succesful Academic Writing*. New: Guilford Press.
- Antonioni, D. & Park, H. (2001). The relationship between penilai affect and three sources of 360 degree feedback ratings. *Journal of Management*, 27, 479-495.
- Alwiyah, Tahir, I. M., & Sayyida. (2016). Persepsi Dosen Terhadap Penelitian. Lanskap Demografi di Indonesia dalam Masyarakat Ekonomi Asean (MEA) (p. 258). Sumenep: Universitas Wiraraja Sumenep.
- Attali, J., Benaissa, A., Soize, S., Kadziolka, K., Portefaix, C., & Pierot, L. (2014). *Follow-up of intracranial aneurysms treated by fl ow diverter : comparison of three-dimensional time-of- flight MR angiography ( 3D-TOF-MRA ) and contrast- enhanced MR angiography ( CE-MRA ) sequences with digital subtraction angiography as the gold stand*. 1–6. <https://doi.org/10.1136/neurintsurg-2014-011449>
- Athey, T. R., & McIntyre, R. M. (1987). Effect of penilai training on penilai accuracy: Levels-ofprocessing theory and social facilitation theory perspectives. *Journal of Applied Psychology*, 72, 567-572.

- Atkinson, J. W. (1981). Studying personality in the context of an advanced motivational psychology. *American Psychologist*, 36(2), 117–128. <https://doi.org/10.1037/0003-066X.36.2.117>
- Baillargeon, J., Penn, J. V, Thomas, C. R., Temple, J. R., Baillargeon, G., & Murray, O. J. (2009). *Psychiatric Disorders and Suicide in the Nation ' s Largest State Prison System*.
- Barkaoui, K. (2013). *Using Multilevel Modeling in Language Assessment Research : A Conceptual Introduction Using Multilevel Modeling in Language Assessment Research : A Conceptual Introduction*. (August), 37–41. <https://doi.org/10.1080/15434303.2013.769546>
- Bartels, L. K., & Doverspike, D. (1997). Assessing the assessor: The relationship of assessor personality to leniency in assessment center ratings. *Journal of Social Behavior and Personality*, 12, 179-190.
- Bernardin, H. J., Orban, J. A., & Carlyle, J. J. (1981). Performance rating as a function of trust in appraisal and penilai individual differences. *Academy of Management Proceedings*, 311-315.
- Bernardin, Cooke & Villanova (2000). Conscientiousness and agreeableness as predictors of rating leniency. *Journal of Applied Psychology*, 85, 232-234.
- Besterfield-Sacre, M., Gerchak, J., Lyons, M. R., Shuman, L. J., & Wolfe, H. (2004). Scoring concept maps: An integrated rubric for assessing engineering education. *Journal of Engineering Education*, 93(2), 105–115. <https://doi.org/10.1002/j.2168-9830.2004.tb00795.x>
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences (2nded.)*. United States of America: Lawrence Erlbaum Associates, Inc.
- Bono, J. E., & Judge, T. A. (2004). Personality and transformational and transactional leadership. *Journal of Applied Psychology*, 89, 901-910.
- Boone, W. J., **Staver**, J. R., & **Yale**, M. S. (2014). *Rasch Analysis in the Human Sciences*. Dordrecht: Springer. <http://dx.doi.org/10.1007/978-94-007-6857-4>.
- Borman, W. C., & Hallam, G. L. (1991). Observation accuracy for assessors of work-sample performance: Consistency across task and individual-differences correlates. *Journal of Applied Psychology*, 76, 11-18.

- Borman, W.C. (1977). Consistency of rating accuracy and rating errors in the judgment of human performance. *Organizational Behavior and Human Performance*, 20, 238-252.
- Borman, W. C. (2004). Introduction to the special issue: Personality and the prediction of job performance: More than the big five. *Human Performance*, 17, 267-269.
- Brookhart, S. M. (2001). *Developing Measurement Theory for Classroom Assessment Purposes and Uses*.
- Burrows, T. (2011). *Writing research articles for publication*. Unpublished manuscript, the Asian Institute of Technology Language Center, Khlong Luang, Thailand.
- Cagnone, S., & Ricci, R. (2005). *Student Ability Assessment Based on Two IRT Models*. 2(2), 209–218.
- Caldwell, D. F., & Burger, J. M. (1997). Personality and social influence strategies in the workplace. *Personality and Social Psychology Bulletin*, 1003-1012.
- Carrell, P. (1995). The effect of writers' personalities and penilais' personalities on the holistic evaluation of writing. *Assessing Writing*, 2(2), 153–190.
- Cleveland, J. N., Murphy, K. R., & Williams, R. E. (1989). Multiple uses of performance appraisal: Prevalence and correlates. *Journal of Applied Psychology*, 74, 130-135.
- Chaer, Abdul. 2011. *Ragam Bahasa Ilmiah*. Jakarta: Rineka Cipta.
- Cumming, A. (1990). *Language Testing Expertise in evaluating second language compositions*. <https://doi.org/10.1177/026553229000700104>
- Dunbar, N. E., Brooks, C. F., & Kubicka-miller, T. (2006). *Oral Communication Skills in Higher Education : Using a Performance-Based Evaluation Rubric to Assess Communication Skills 1*. 31(2), 115–128. <https://doi.org/10.1007/s10755-006-9012-x>
- Dunsmuir, Sandra & Clifford, Vivienne. (2003). Children's Writing and the Use of Information and Communications Technology. *Educational Psychology in Practice*. 19. 170-187. 10.1080/0266736032000109447.
- Eckes, T. (2008). *Language Testing*. <https://doi.org/10.1177/0265532207086780>
- Eckes, T. (2011). *Introduction to Many-Facet Rasch Measurement*.

- Elsasser, D. A., Ananth, C. V, Prasad, V., & Vintzileos, A. M. (2010). *European Journal of Obstetrics & Gynecology and Reproductive Biology Diagnosis of placental abruption: relationship between clinical and histopathological findings*. 148, 125–130. <https://doi.org/10.1016/j.ejogrb.2009.10.005>
- Engelhard, G. (2012). *Examining Rating Quality in Writing Assessment : Penilai Agreement , Error , and Accuracy*. 13(4).
- Engelhard, G. Jr. (1994). Examining penilai errors in the assessment of written composition with a many-faceted Rasch model. *Journal of Educational Measurement* 31/2: 93–112.
- Elder, C., Barkhuizen, G., Knoch, U., & Randow, J. (2007). Evaluating penilai responses to an online training program for L2 writing assessment. *Language Testing*, 24 (1), 37-64.
- Farh, J. L., & Dobbins, G. H. (1989). Effects of self-esteem on leniency bias in self-reports of performance: A structural equation model analysis. *Personnel Psychology*, 42, 835-850.
- Farrokhi, F., Esfandiari, R., & Schaefer, E. (2012). *A Many-Facet Rasch Measurement of Differential Penilai Severitas / Leniency and Teacher Assessment*. 2(9), 8786–8798.
- Feist, Gregory J., Feist J., & T.A. Roberts (2021). *Theories of personality*. Nwe York: McGraw-Hill
- Freedman, S. W. (1981). Influences on evaluators of expository essays: Beyond the text. *Research in the Teaching of English*, 15 (3), 245-55.
- Ferne, T., & Rupp, A. A. (n.d.). *A Synthesis of 15 Years of Research on DIF in Language Testing : Methodological Advances , Challenges , and Recommendations*. (July 2013), 37–41.
- Forret, M. L., & Dougherty, T. W. (2001). Correlates of networking behavior for managerial and
- Fred Pirczak. (2017). *Evaluating Research in Academic Journals* (Sixth). New York: Routledge.
- Giuliani, A., Fernandez, M., Farinelli, M., Baratto, L., Capra, R., Rovetta, G., ... Calzà, L. (2004). Guidelines on writing a first quantitative academic article. *International Journal of Tissue Reactions*, 26(1–2), 29–37. <https://doi.org/10.1073/pnas.0703993104>

- Hamp-Lyons, L, 1991. Scoring procedures in ESL contexts. In: Assessing second language writing in academic context (ed L. Hamp-Lyons.) pp.241-276. Norwood NJ, Ablex.
- Hamp-Lyons, L, 1995. Rating nonnative writing: The trouble with holistic scoring. *TESOL Quarterly*, 29(4): 759-762.
- Hamp-Lyons, L, 2003. Writing teachers as assessors of writing. In: Exploring the dynamics of second language writing. (Ed B. Kroll) pp.162-189. Cambridge, Cambridge University Press.
- Harari, Michael & Rudolph, Cort & Laginess, Andrew. (2014). Does rater personality matter? A meta-analysis of penilai Big Five–performance rating relationships. *Journal of Occupational and Organizational Psychology*. 88. 10.1111/joop.12086.
- Hogue, A., Henderson, C. E., Dauber, S., Barajas, P. C., Fried, A., & Liddle, H. A. (2008). *Treatment Adherence , Competence , and Outcome in Individual and Family Therapy for Adolescent Behavior Problems*. 76(4), 544–555. <https://doi.org/10.1037/0022-006X.76.4.544>
- Huang, J., & Foote, C. J. (2010). *Grading Between the Lines : What Really Impacts Professors ' Holistic Evaluation of ESL Graduate Student Writing ? Grading Between the Lines : What Really Impacts Professors ' Holistic Evaluation of ESL Graduate Student Writing ?* (September 2013), 37–41. <https://doi.org/10.1080/15434300903540894>
- Islam, U., Ulum, D., & Lamongan, U. (n.d.). *ARTIKEL ILMIAH : ANATOMI , BAHASA , DAN KESALAHANNYA*. 1–10.
- Jalali, S., Seifbarghy, M., Sadeghi, J., & Ahmadi, S. (2016). *Knowledge-Based Systems Optimizing a bi-objective reliable facility location problem with adapted stochastic measures using tuned-parameter multi-objective algorithms*. 95, 45–57. <https://doi.org/10.1016/j.knosys.2015.12.001>
- James Hartley. (2008). *Academic Writing and Publishing*. New York: Taylor & Francis e-Library.
- Jawahar & Stone (1997) Influence of penilais' self-consciousness and appraisal purpose on leniency and accuracy of performance ratings. *Psychological Reports*, 80, 323-337.
- Jawahar, I. M. (2001). Attitudes, self-monitoring, and appraisal behaviors. *Journal of Applied Psychology*, 86, 875-883.

- Jelley, R. B. (2005). Penilai personality and performance-evaluation leniency. Unpublished doctoral dissertation, University of Western Ontario, Canada.
- Jenkins, M., & Griffith, R. (2004). Using personality constructs to predict performance: Narrow or broad bandwidth. *Journal of Business and Psychology*, *19*, 255-269.
- Jensen-Campbell, L. A., & Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal
- Johanna & Brydges, Bruce. (2006). Improving the Quality of University Libraries through Citation Mining and Analysis Using Two New Dissertation Bibliometric Assessment Tools. Online Submission.
- Jonsson, A., & Svingby, G. (2007). *The use of scoring rubrics : Reliability , validity and educational consequences.* *2*, 130–144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, *87*, 530-541.
- Kane, J. S. (1994). A model of volitional rating behavior. *Human Resource Management Review*, *4*, 283-310.
- Kane, J. S., Bernardin, H. J., Villanova, P., & Peyrefitte, J. (1995). Stability of penilai leniency: Three studies. *Academy of Management Journal*, *38*, 1036-1051.
- Kingstrom, P. O., & Bass, A. R. (1981). A critical analysis of studies comparing behaviorally anchored rating scales (BARS) and other rating formats. *Personnel Psychology*, *34*, 263-289.
- Knoch, U., Read, J., & Randow, J. Von. (2007). *Re-training writing penilais online : How does it compare with face-to-face training?* *12*, 26–43. <https://doi.org/10.1016/j.asw.2007.04.001>
- Kondo-Brown, K. (2002). A FACETS analysis of penilai bias in measuring Japanese second language writing performance. *Language Testing*, *19*(1), 3-31.
- Kotzé, T. (2007). *Guidelines on writing a first quantitative academic article.*
- Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin*, *87*, 72-107. Lee, Ashton & Shin (2005) Personality correlates of workplace anti-social behavior. *Applied Psychology: An International Review*, *54*, 81-98.

- LePine, J. A., Erez, A., & Johnson, D. E. (2002). The nature and dimensionality of organizational citizenship behavior: A critical review of meta-analysis. *Journal of Applied Psychology*, 87, 52-65.
- Linacre J. (2004). Rasch model estimation: further topics. *Journal of Applied Measurement*, 5(1), 95–110.
- Linden, W. J. Van Der, & Hambleton, R. K. (1997). *Item Response Theory : Brief History , Common Models , and Extensions*. 1–2.
- Linacre, J. M. (1989). Many faceted Rasch measurement. Chicago: MESA Press.
- Linacre, J. M. (2013). A user's guide to Facets: Rasch measurement computer program [Computer program manual]. Chicago: MESA Press.
- Linacre, J. M. (2013). A user's guide to Facets: Rasch measurement computer program [Computer program manual]. Chicago: MESA Press
- Linacre, J. M. (2015). FACETS [Computer program, version 3.71.4]. Chicago: MESA Press.
- Nunally & Bernstein. (1994). *Psychometric theory*. New York: McGraw Hill
- Popham, W. J. (1990). *Modern educational measurement: a practitioner's perspective* .
- Lumley, T. & McNamara, T. F. (1995). Penilai characteristics and penilai bias: Implications for training. *Language Testing* 12(1), 55-71.
- Lumley, T. & McNamara, T. F. (1995). Penilai characteristics and penilai bias: Implications for training. *Language Testing*, 12, 54-71.
- Lumley, T. (2002). Assessment criteria in a large-scale writing test: What do they really mean to the penilais? *Language Testing*, 19(3): 246–276.
- Lumley, T. (2005). *Assessing second language writing: The penilai's perspective*. Frankfurt am Main: Peter Lang.
- Lynch, B. K., & McNamara, T. F. (1998). Using G-theory and Many-facet Rasch measurement in the development of performance assessments of the ESL speaking skills of immigrants. *Language Testing*, 15(2), 158–180. <https://doi.org/10.1177/026553229801500202>
- Mandell (1956) Supervisory characteristics and ratings: A summary of recent research. *Personnel Psychology*, 32, 435-440.
- Mardapi, D. (2004). *Penyusunan Tes Hasil Belajar*. Yogyakarta: Program Pascasarjana Universitas Negeri Yogyakarta.

- Mardapi, D. (2012). *Pengukuran, Penilaian dan Evaluasi Pendidikan*. Yogyakarta: Nuha Litera.
- Markhamah. 2000. "Bahasa dalam Penulisan Ilmiah". Dalam Prayitno, Harun Joko dkk. (ed.). 2000. *Pembudayaan Penulisan Karya Ilmiah*. Surakarta: Muhammadiyah University Press.
- Maryadi. 2000. "Pengertian dan Kriteria Karya Ilmiah". Dalam Prayitno, Harun Joko dkk. (ed.). 2000. *Pembudayaan Penulisan Karya Ilmiah*. Surakarta: Muhammadiyah University Press.
- Masters, G. N. (1982). *A Rasch Modeln For Partial Credit Scoring*. 47(2), 149–150.
- Matsuno, S, 2009. Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26(1): 75–100.
- Mcnamara, D. S., Kintsch, E., Songer, N. B., Mcnamara, D. S., & Kintsch, E. (2009). *Are Good Texts Always Better? Interactions of Text Coherence , Background Knowledge , and Levels of Understanding in Learning From Text Are Good Texts Always Better? Interactions of Text Coherence , Background Knowledge , and Levels of Understanding in.* (January 2015), 37–41. <https://doi.org/10.1207/s1532690xci1401>
- Mcnamara, T., & Knoch, U. (2012). *Language Testing The Rasch wars : The*. <https://doi.org/10.1177/0265532211430367>
- McNamara, T. F. (1996). *Measuring second language performance*. London: Longman.
- Mei, W. S. (2001). *Investigating Penilais ' Use of Analytic Descriptors in Assessing Writing*. 9(2), 69–104.
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. Mertler, Craig A. *Research & Evaluation*, 7(25), 1–10. Retrieved from <http://pareonline.net/getvn.asp?v=7&n=25><http://PAREonline.net/getvn.asp?v=7&n=25>.
- Mertler, Craig A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). Available online: <http://PAREonline.net/getvn.asp?v=7&n=25>.
- Messics, S. (1996). Validity and wahsback in language testing. *Language testing*, 241-256. Meyer, H. H. (1980). Self-appraisal of job performance. *Personnel Psychology*, 33, 291-295.



- Meyers, A. (2006). *Composing with confidence: Writing effective paragraphs and essays*. New York, Pearson/Longman.
- Miller, J. D., & Lynam, D. (2001). Structural models of personality and their relation to antisocial behavior: A meta-analytic review. *Criminology*, *39*, 765-795.
- Mostafaei Alaei, Mahnaz & Ahmadi, Masoume & Sabourian Zadeh, Naser. (2014). The Impact of Penilai's Personality Traits on Holistic and Analytic Scores: Does Genre Make any Difference too?. *Procedia - Social and Behavioral Sciences*. 98. 10.1016/j.sbspro.2014.03.539.
- Moskal, B. M. (2000). Scoring rubrics: what, when and how?. Moskal, Barbara M. *Practical Assessment, Research & Evaluation*, *7*(3), 3-7. Retrieved from <http://pareonline.net/getvn.asp?v=7&n=3>
- Mount, M. K. (1984). Psychometric properties of subordinate ratings of managerial performance. *Personnel Psychology*, *37*, 687-702.
- Mount, M. K., & Barrick, M. R. (1998). Five reasons why the "Big Five" article has been frequently cited. *Personnel Psychology*, *51*, 849-857.
- Myford, C. M., & Wolfe, E. W. (2004). Detecting and Measuring Penilai Effects Using Many-Facet Rasch Measurement: Part I. *Journal of Applied Measurement*, *5*(2), 189-227.
- Murphy, K. R., & Balzer, W. K. (1989). Penilai errors and rating accuracy. *Journal of Applied Psychology*, *74*, 619-624..
- Myford, C., & Wolfe, E. W. (2004). Detecting and Measuring Penilai Effects Using Many-Facet Rasch Measurement: Part II. *Journal of Applied Measurement*, *5*(2), 189-227.
- Nanang, Oleh & Subekti, Nanang. (2018). *Rangking Publikasi Ilmiah Internasional Indonesia*. 10.13140/RG.2.2.20404.96641.
- Norris, J. M., Brown, J. D., Hudson, T. D., Bonk, W., Norris, J. M., Arizona, N., & James, D. (2002). *Language Testing language performance assessment performance assessment*. <https://doi.org/10.1191/0265532202lt237oa>
- Nugraha, A. J., Suyitno, H., & Susilaningsih, E. (2017). *Analisis Kemampuan Berpikir Kritis Ditinjau dari Keterampilan Proses Sains dan Motivasi Belajar melalui Model PBL Abstrak*. *6*(1), 35-43.

- Nunnally, J. C. (1978). *Psychometric theory*. (2nd ed.) New York, NY: McGraw-Hill.
- Paunonen, S. V. (1998). Hierarchical organization of personality and prediction of behavior. *Journal of Personality and Social Psychology*, 74, 538-556.
- Osterlind, S. J. (1983). Test item bias. Beverly Hills, CA: Sage Publication
- Pasteur, U. L. (1996). *Organization Industrial Horizontal and vertical differentiation : The Launhardt model*. 14, 485–506.
- Paul J. Silvia, P. (1976). *Up*. APA LifeTools.
- Paunonen, S. V., Rothstein, M. G., & Jackson, D. N. (1999). Narrow reasoning about the use of broad personality measures for personnel selection. *Journal of Organizational Behavior*, 20, 389-405.
- Perry, J. M., He, X. C., Sugimura, R., & Li, L. (2013). *Stem Cell Dormancy : Maintaining a Reserved Population Keywords*. 1–14.
- Primi, R., & Silvia, P. J. (2018). 3 4 4 1. professional employees. *Group and Organization Management*, 26, 283–311. doi:10.1177/1059601101263004
- Reddy, Y. M., & Andrade, H. (n.d.). *Assessment & Evaluation in Higher Education A review of rubric use in higher education*. (May 2013), 37–41. <https://doi.org/10.1080/02602930902862859>
- Reddy, Y., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation In Higher Education*, 35, 435- 448. doi:10.1080/02602930902862859
- Rooks, Terry & T. Kane, Michael & Cohen, Allan. (1996). Threats to the Valid Use of Assessments. *Assessment in Education: Principles, Policy & Practice*. 3. 265-286. 10.1080/0969594960030302.
- Rufaidah VW. 2010. Produktifitas Publikasi Peneliti Badan Litbang Pertanian. *Jurnal Perpustakaan Pertanian*. Vol. 19, Nomor 1. Saal, F. E., Downey, R. G., & Lahey, M. A. (1980). Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin*, 88, 413-428.
- Saito, H, 2008. EFL classroom peer assessment: Training effects on rating and commenting. *Language Testing*, 25(4): 553-581.
- Saito, H. and T. Fujita, 2004. Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching Research*, 8(1): 31-54.

- Saito, H. and T. Fujita, 2009. Peer-assessing peers' contribution to EFL group presentations. *RELC Journal*, 40(2): 149- 171.
- Salovey, P., Rothman, A. J., Detweiler, J. B., Steward, W. T., & Universi, Y. (2000). *Emotional States and Physical Health*. 55(I), 110–121.
- Schaefer, E, 2008. Penilai bias patterns in an EFL writing assessment. *Language Testing*, 25(4): 465–493.
- Schieritz, N., & Größler, A. (2002). *Emergent Structures in Supply Chains — A Study Integrating Agent-Based and System Dynamics Modeling*.
- Sebok, S. S., Luu, K., & Klinger, D. A. (2014). *Psychometric properties of the multiple mini-interview used for medical admissions: findings from generalizability and Rasch analyses*. 71–84. <https://doi.org/10.1007/s10459-013-9463-7>
- Shohamy, E., Gordon, C. M., & Kramer, R. (1992). The effects of penilais' backgrounds and training on the reliability of direct writing tests. *Modern Language Journal*, 76 (1), 27-33.
- Simon, M., & Forgette-Giroux, R. (2001). *A rubric for scoring postsecondary academic skills - Practical Assessment, Research & Evaluation*. 7(18), 1–7.
- Sumintono, B. (2014). Model Rasch untuk Penelitian Sosial Kuantitatif. *Makalah Kuliah Umum Di Jurusan Statistika, ITS Surabaya, 21 November 2014*, (November 201), 1–9. <https://doi.org/10.1002/rcm.1134>
- Sumintono dan Wahyu Widhiarso. (2013). *Aplikasi Model Rasch untuk Penelitian Ilmu-Ilmu Sosial*. Jakarta: TrimKom
- Suwito. 1982. *Pengantar Awal Sociolinguistik: Teori dan Problema*. Surakarta: Henary Offset
- Syarif, M. (n.d.). *Teori dan Model Pengembangan Kelembagaan Pendidikan Tinggi Islam*. 28(3).
- Tagger, S., & Brown, T. C. (2006). Interpersonal affect and peer rating bias in teams. *Small Group Research*, 37, 86-111
- Tcherni-buzzeo, M. (n.d.). *Evaluating Research in Academic Journals*.
- Tsui, A. S., & Barry, B. (1986). Interpersonal affect and rating errors. *Academy of Management Journal*, 29, 586-599.

- Waldman & Thornton (1988) A field study of rating conditions and leniency in performance appraisal. *Psychological Reports*, 63, 835-840.
- Weigle, S. C. (1994b). *Effects of training on penilais of English as a second language. compositions: Qualitative and quantitative approaches*. Unpublished Ph.D dissertation, University of California, Los Angeles.
- Weigle, S. C, 2002. *Assessing Writing*. Cambridge: Cambridge University Press.
- Wen, L., & Chase, C. D. (1999). *Pleiotropic effects of a nuclear restorer-of-fertility locus on mitochondrial transcripts in male-fertile and S male-sterile maize*. 1997–1998.
- Widhiarso, W. (2010). *Melibatkan Penilai dalam Pengembangan Alat Ukur*. 1–4.
- Wibowo, Wahyu. 2008. *Piawai Menembus Jurnal Terakreditasi: Paradigma Baru Kiat Menulis Artikel Ilmiah*. Jakarta: PT Bumi Aksara.
- Wigglesworth, G. (1993). Exploring bias analysis as a tool for improving penilai consistency in assessing oral interaction. *Language Testing*, 10, 305-23.
- Winke, P., S., Gass, and C. Myford, 2011. *The relationship between penilais' prior language study and the evaluation of foreign language speech samples (TOEFL iBT Research Report TOEFL iBT-160)*. Princeton, NJ, Educational TestingService.