

## REFERENCES

- Abbott, M. L. (2006). ESL Reading Strategies: Differences in Arabic and Mandarin Speaker Test Performance. *Language Learning. A Journal of Research in Language Studies*, Vol. 56 (4) Pg. 633-670. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9922.2006.00391.x/pdf>.
- Alderson , J. C. (2000). *Assessing Reading* . Cambridge : Cambridge University Press.
- Alderson , J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Alderson, J. C. (2005). *Diagnosing Foreign Language Proficiency: The Interface between Learning and Assessment*. London : Continuum.
- Arikunto, S. (2005). *Dasar - Dasar Evaluasi Pendidikan* . Jakarta : Bumi aksara.
- Aulia, I. F., Sukirlan, M., & Sudirman. (2014). *Analysis of the Quality of Teacher Made-Reading Comprehension Test Items Using ITEMAN*. Bandar Lampung. Retrieved from <https://media.neliti.com/media/publications/194356-EN-analysis-of-the-quality-of-teacher-made.pdf>.: University of Lampung .
- Bachman, L. F., & Palmer, A. S. (1996). *Language Testing in Practice: Designing and Deveoping Useful Language Test*. Walton Street, Oxford: Oxford University Press.
- Bailey, E. (2015). *Reading Comprehension Skills: Making Predictions* . Retrieved at August 11th 2017 from <http://specialed.about.com/od/readingliteracy/a/Reading-Comprehension-Skills-Making-Predictions.htm>.
- Berg, B. L. (2001). *Qualitative Research Methods for the Social Sciences*. USA: Allyn & Bacon.
- Blomeke, S., & Delaney, S. (2012). Assessment of teacher knowledge across countries: a review of the state of research. *Springer*, Vol. 4 (3) Retrieved at 13 July 2017 from <http://www.springerlink.com/content/w7214t0124j75835/>.
- Brown , H. D. (2004). *Language Asseement. Principles and Classroom Practices*. New York: Addison Wesley Longman, Inc.
- Burke, K. (1999). *How to Assess Authentic Learning , 3rd Edition*. London: Pearson SkyLight Training and Publishing Inc.

- Caena, F. (2011). Literature Review. Teacher's Core Competencies: Requirements and Development. *European Commission*, Retrieved at 13 July 2017 from [http://ec.europa.eu/dgs/education\\_culture/repository/education/policy/strategic-framework/doc/teacher-competences\\_en.pdf](http://ec.europa.eu/dgs/education_culture/repository/education/policy/strategic-framework/doc/teacher-competences_en.pdf).
- Cheek, Jr., E., & Collins, M. (1985). *Strategies for Reading Success*. Columbus, Ohio : Charles E. Merrill Publishing Company.
- Cheung, D., & Bucat, R. (2002). How Can We Construct Good Multiple Choice Items? *Hongkong: Science and Technology Education Conference*.
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education 7th (ed)*. USA: Routledge .
- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluation Quantitative and Qualitative Research*. Boston: Pearson Education.
- Croker, R., & Heigham, J. (2009). *Qualitative Research in Applied Linguistics. A Practical Introduction*. Great Britain: Palgrave Macmillan.
- Crooks, T. J. (1988). The Impact of Classroom Evaluation Practices on Students . *Sage Online Journal* , Vol. 58 (4), p. 438-481.
- Day, R. R., & Bamford, J. (1998). *Extensive Reading in The Second Language Classroom* . Cambridge : Cambridge University Press.
- Day, R. R., & Park, J.-s. (2005). Developing Reading Comprehension Questions . *Reading in a Foreign Language* , Vol. 17(1).
- Disha. (2012). *The Next Generation Library*. Retrieved from Teacher Made-Test: Meaning, Features, and Uses: <http://www.yourarticlelibrary.com/statistics-2/teacher-made-test-meaning-features-and-uses-statistics/92607>
- Flateby, T. L. (2013). *A Guide for Writing and Improving Achievement Tests*. South Florida: USF Press. Retrieved from <http://www.usf.edu/provost/documents/assessment/resources-guideforwritingandimprovingachievementtests.pdf>.
- Frey, B. B. (2012). *KU. The university of Kansas*. Retrieved from The quality test construction : [http://www.specialconnections.ku.edu/?q=assessment/quality\\_test\\_construction](http://www.specialconnections.ku.edu/?q=assessment/quality_test_construction)
- Gisev, N., Bell, J. S., & Chen, T. F. (2013). Inter-rater agreement and inter-rater reliability: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*. , 9, 330-338.
- Grabe, W., & Stoller, F. L. (2002). *Reading and Researching Reading*. London: Pearson Education. Retrieved 1st March 2017 from <http://nflrc.hawaii.edu/rfl/October2002/reviews/hood.pdf>.

- Grellet, F. (1999). *Developing Reading Skills: A Practical Guide to Reading Comprehension*. Cambridge: Cambridge University Press. Retrieved at August 7th 2017 from <http://celta.wikispaces.com/file/view/Grellet%20Chapter%201.pdf>.
- Grellet, F. (2004). *Developing Reading Skills: A Practical Guide to Reading Comprehension Exercises*. Cambridge : Cambridge University Press. Retrieved at 9th August 2017 from [http://assets.cambridge.org/97805212/83649/excerpt/9780521283649\\_excerpt.pdf](http://assets.cambridge.org/97805212/83649/excerpt/9780521283649_excerpt.pdf).
- Hanson, M. (1996). *Self-Management through Self-Monitoring*. In K. Jones & T. Charlton (Eds). *Overcoming Learning and Behaviour Difficulties: Partnership with Pupils*. London: Routledge.
- Harmer , J. (2001). *The Practice of English Language Teaching 3rd Edition*. London: Longman.
- Harmer, J. (1991). *The Practice of English Language Teaching*. London: Longman Group UK Limited.
- Harmer, J. (2007). *How to Teach English* . Oxford : Pearson Education Limited .
- Hill, C., & Parry, K. (1994). *Models of Literacy: the nature of reading tests*. From *Testing to Assessment: English as an International language*. Harlow, UK: Longman.
- Hopkins, K. D., Stanley, J. C., & Hopkins, B. R. (1990). *Educational and psychological measurement and evaluation*. Needham Height, MA: Allyn and Bacon.
- How to Judge the Quality of an Objective Classroom Test. (2006). *Evaluation and Examination Service. The University of Iowa*, Retrieved from [http://people.uncw.edu/caropresoe/MCCSSS%20Workshop/Online%20Files/UIA\\_How-to-judge-quality-of-test\\_2006.pdf](http://people.uncw.edu/caropresoe/MCCSSS%20Workshop/Online%20Files/UIA_How-to-judge-quality-of-test_2006.pdf).
- Hughes , A. (2003). *Testing for Language Teachers*. Cambridge : Cambridge University Press.
- IELTS-Up. (2017). *IELTS Reading Lesson 5*. Retrieved at 29 September 2017 from <http://ielts-up.com/reading/short-answer-questions-lesson.html>.
- King, N., & Horrocks, C. (2012). *Interviews in Qualitative Research*. London: SAGE Publication Inc.
- Latief. (2011, May Monday). Menilai Siswa, Tak Cukup Hanya Tes.
- Lei, L. (2008). Validation of the C-test amongst Chinese ESL Learners. *The Journal of ASIA TEFL*, Vol. 5(2) pp.117-140. Retrieved at 29th September 2017 from [http://www.asiatefl.org/main/download\\_pdf.php?i=266&c=1419312234](http://www.asiatefl.org/main/download_pdf.php?i=266&c=1419312234) .

- Lynch, M. (2016, June 28). *Education Week*. Retrieved from Education Futures: Emerging Trends in K-12: [http://blogs.edweek.org/edweek/education\\_futures/2016/06/three\\_important\\_critiques\\_of\\_standardized\\_assessments.html](http://blogs.edweek.org/edweek/education_futures/2016/06/three_important_critiques_of_standardized_assessments.html)
- Mahoney, S. (2009, May 26). *Common Rating Scales to Use When Writing Questions*. Retrieved from Verint Customer Engagement: Retrieved at 20 Nov 2017 from <http://blog.verint.com/customer-engagement/common-rating-scales-to-use-when-writing-questions>
- Malik, R. S., & Hamied, F. A. (2014). *Research Methods: A Guide for First Time Researchers*. Bandung: UPI Press.
- Marmah, A. A., & Impraim, A. K. (2013). University Lecturer's Competence in the Construction of Multiple Choice Test Items: A Case Study of Coltek-Kumasi. *The International Journal of Humanities and Social Studies*, Vol. 4 (1), p. 1-9. Retrieved at 10 January 2017 from <http://theijhss.com/october2013/1.HSS1310-002.pdf>.
- Marso, R. N., & Pigge, F. L. (1988). An Analysis of teacher-made test and testing: classroom resources, guidelines, and practices. *Paper presented at the annual meeting of the Mid-Western Research Association, Chicago*.
- Mart, C. T. (2012). Guessing the Meaning of Words from Context: Why and How. *International Journal of Applied Linguistics and English Literature*, Vol. 1(6). Retrieved at 12 September 2017 from <http://www.journals.aiac.org.au/index.php/IJALEL/article/viewFile/819/751>.
- Maxwell, M. J. (1972). Skimming and Scanning Improvement: The Needs, Assumptions, and Knowledge Base. *Sage Journal*, Vol. 5(1).
- Mertler, C. A. (2009). Teachers' Assessment Knowledge and Their Perceptions of the Impact of Classroom Assessment Professional Development. *Sage Journal*, Vol. 12(2) p. 101-113. Retrieved at 1 March 2017 from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.921.5783&rep=rep1&type=pdf>.
- MR, F. (2014). The Comparison of Two Test Item Format in Assessing Students' Reading Comprehension Ability. *ELTIN Journal*, Vol.2(2). Retrieved at 28th June 2017 from <http://e-journal.stkipsiliwangi.ac.id/index.php/eltin/article/view/49/48>.
- Musthafa, B. (1996). Assessment Strategies for Optimal Literacy. *Educational Resources Information Center (ERIC)*, Retrieved at 18 May 2017 from <http://files.eric.ed.gov/fulltext/ED403300.pdf>.
- Nation, I. (2008). *TEaching ESL/EFL Reading and Writing. ESL & Applied Linguistics professional Series*. New York: Routledge.

- Ong, S. L. (2011). *Profiling Classroom Teachers Assessment* . Oxford. (2006). *Word Power Dictionary*. New York: Oxford University Press.
- Parr, A. M., & Bauer, W. (2006). *Teacher made tes reliability: A comparison of test score and students study habits from Friday to Monday in a high school biology class in Monroe County, Ohio*. Ohio: Marietta College.
- Patton, M. (1990). *Qualitative Evaluation and Research Methods (2nd edition)*. Newbury Park, CA: Sage.
- Razali , K., & Jannah , M. (2015). The Comparison Between National Final Examination Test Items and English Teacher Made-Test Items of 2010-2011. Vol. 22 (1) p.10-22.
- Reiner, C. M., Bothell, T. W., Sudweekw, R. R., & Wood, B. (2002). *Preparing effective essay questions. a self-directed workbook for educators*. Birmingham: New Forum Press.
- Richards , J. C. (1990). *The Language Teaching Matrix*. Cambridge : Cambridge University Press.
- Richards , J. C. (2011). Exploring Teacher Competence in Language Teaching . *JALT* , 3-7.
- Rock, M. (2005). Use of Strategic Self-Monitoring to Enhance Academic Engagement, Productivity, and Accuracy of Students With and Without Exceptionalities. *Sage Journals* , Vol. 7(1). Retrieved at 12 July 2017 from <http://journals.sagepub.com/doi/pdf/10.1177/10983007050070010201> .
- Saefurrohman , & Balinas , E. S. (2016). English Teachers Classroom Assessment Practices . *International Journal of Evaluation and Research in Education (IJERE)* , Vol. 5(1), p. 82-92.
- Sah, K. P. (2012). Assessment and Test in Teaching and Learning. *Academic Voices, A Multidisciplinary Journal*, Vol. 2(1). Retrieved at 22nd August 2017 from <https://www.google.co.id/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&uact=8&ved=0ahUKEwinn7yUr7fWAhVFrI8KHdz9BVAQFg hFMAQ&url=https%3A%2F%2Fwww.nepjol.info%2Findex.php%2FAV%2Farticle%2FviewFile%2F8282%2F67>.
- Shweta, Bajpai, R. C., & Chaturvedi, H. K. (2015). Evaluation of Inter-Rater Aggreemeent and Inter-Rater Reliability for Observational Data: AN Overview of Concepts and Methods . *Journal of the Indian Academy of Applied Psychology*, Vol. 41No. 3 (Special Issue), 20-27.
- Souhila, R. (2014). The Use of Reading Strategies in Improving Reading Comprehension. *Dissertation* .

- Suah , S. L., & Ong, S. L. (2012). Investigating Assessment Practices of In-Service Teachers. *International Online Journal of Educational Sciences* , Vol. 4(1) p. 91-106. Retrieved at 15 February 2017 from [http://www.iojes.net/userfiles/Article/IOJES\\_724.pdf](http://www.iojes.net/userfiles/Article/IOJES_724.pdf).
- Suherdi, D. (2013). *Buku Pedoman Penyelenggaraan Pendidikan Profesi Guru Bahasa Inggris. Buku Ajar Pemantapan Kompetensi Akademik 3.1*. Bandung : Celtic Press.
- Thornbury, S. (2002). *How to Teach Vocabulary*. Harlow: Longman.
- Tshabalala, T., Mapolisa, T., Gazimbe, P., & Ncube, A. C. (2015). Establishing the Effectiveness of Teacher-Made Tests in Nkayi District Primary Schools. *Nova Journal of Humanities and Social Sciences* , Vol. 4(1) p. 1-6. Retrieved at 3 March 2017 from [www.novaexplore.com](http://www.novaexplore.com).
- Wallace, C. (1992). *Reading*. Walton Sreet, Oxford: Oxford University Press.
- Worthen , B. R., Borg , W. R., & White , K. R. (1993). Journal of Psychoeducational Assessment. *MEasurement and Evaluations in the School*, 13, 1995, p. 416-418. Retrieved 27 February 2017 from <http://journals.sagepub.com/doi/abs/10.1177/073428299501300409?journalCode=jpaa>.
- Zhang , Z., & Burry-Stock, J. A. (2003). Classroom Assessment Practices and Teachers' Self-Perceived Assessment Skills. *Applied Measurement in Education* , Vol. 16(4) p. 323-342. Retrieved at 15 February 2017 from <http://www.esf.edu/assessment/documents/assessmentpracticesandskills.pdf>.
- Zhang, Z. (1996). Teacher Assessment Competency; A Rasch Model Analysis. *Annual Meeting of American Educational Research Association* (pp. 1-33). New York: Educational Research Information Center.
- Zhao, Z. (2013). An Overview of Studies on Diagnostic Testing and Its Implications for the Development of Diagnostic Speaking Test. *International Journal of English Linguistics*, Vol.3(1). Retrieved at 26th July 2017 from <http://www.ccsenet.org/journal/index.php/ijel/article/download/24239/15353>.