

บรรณานุกรม

- คณะกรรมการมาตรฐานการอุดมศึกษา. (2559). นโยบายการยกระดับมาตรฐานภาษาอังกฤษ
ในสถาบันอุดมศึกษา พ.ศ. 2559. สืบค้นเมื่อวันที่ 28 เมษายน 2567 จาก
http://www.dqe.mhesi.go.th/front_home/Data%20Bhes_2559/04052559.pdf
- ชื่นจิตต์ อธิวรกุล และ จิรดา วุฒนยากร. (2561). การเทียบผลคะแนนแบบทดสอบสมิทธิภาพ
ภาษาอังกฤษ แห่งมหาวิทยาลัยศรีนครินทรวิโรฒกับกรอบมาตรฐานอ้างอิงสากลของยุโรป.
วารสารเทคโนโลยีสุรนารี, 12(2), 69–84.
- พฤต ไบระหมาน. (2559). การประเมินความสอดคล้องระหว่างข้อสอบระดับชาติวิชาภาษาฝรั่งเศส
กับกรอบมาตรฐานความสามารถทางภาษาของสหภาพยุโรป โดยใช้วิธีแผนที่ข้อสอบ
(วิทยานิพนธ์ศึกษาศาสตรมหาบัณฑิต, มหาวิทยาลัยสุโขทัยธรรมาธิราช).
<https://dric.nrct.go.th/Search/SearchDetail/294837>
- มหาวิทยาลัยสุโขทัยธรรมาธิราช. (2560, 14 มิถุนายน). ประกาศมหาวิทยาลัยสุโขทัยธรรมาธิราช เรื่อง
การยกระดับมาตรฐานความรู้ภาษาอังกฤษสำหรับนักศึกษาในระดับปริญญาโท.
- มหาวิทยาลัยสุโขทัยธรรมาธิราช. (2561, 3 กันยายน). ประกาศมหาวิทยาลัยสุโขทัยธรรมาธิราช เรื่อง
การกำหนดเกณฑ์มาตรฐานความรู้ภาษาอังกฤษระดับปริญญาเอก พ.ศ. 2561.
- มหาวิทยาลัยสุโขทัยธรรมาธิราช. (2562, 6 มีนาคม). ประกาศมหาวิทยาลัยสุโขทัยธรรมาธิราช เรื่อง
การยกระดับมาตรฐานภาษาอังกฤษสำหรับนักศึกษาในระดับปริญญาตรี.
- มหาวิทยาลัยสุโขทัยธรรมาธิราช. (2565). โครงการยกระดับมาตรฐานภาษาอังกฤษนักศึกษา มสธ.
สืบค้นเมื่อวันที่ 1 กันยายน 2565 จาก <https://www.stou.ac.th/main/seep.html>
- ศจี จิระโร และ สุภมาส อังสุโชติ. (2563). การตรวจสอบคุณภาพข้อสอบวัดสมิทธิภาพในการใช้
ภาษาอังกฤษของมหาวิทยาลัยสุโขทัยธรรมาธิราช ด้วยระบบคอมพิวเตอร์ตามกรอบการอ้างอิง
ทางภาษาของสหภาพยุโรป (รายงานการวิจัย). นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- ศิริชัย กาญจนวาสี. (2563). การวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ. วารสารวิจัย
การศึกษา, 12(1), 45-60.
- ศุภกรรัตน์ อิงชาติเจริญ. (2567). การตรวจสอบคุณสมบัติทางจิตมิติของข้อสอบวัดความรู้ภาษาอังกฤษ
ของมหาวิทยาลัยสุโขทัยธรรมาธิราช. วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม,
30(1), 214–227.
- สังวรณัฏฐ์ รัตตะโทก. (2561). การวัดและประเมินผลสำหรับการจัดการศึกษาอิงมาตรฐาน. นนทบุรี:
จุฬารัตน์.

- สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน. (2557). *แนวปฏิบัติตามประกาศกระทรวงศึกษาธิการ เรื่อง นโยบายการปฏิรูปการเรียนการสอนภาษาอังกฤษ*. กรุงเทพฯ: บริษัท จามจุรีโปรดักส์ จำกัด.
- สำนักงานคณะกรรมการการอุดมศึกษา. (2559). *นโยบายการยกระดับมาตรฐานภาษาอังกฤษในสถาบันอุดมศึกษา*. สืบค้นเมื่อวันที่ 26 สิงหาคม 2566 จาก <http://surl.li/cseah>
- สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช. (2565). *เกณฑ์การแปลผลคะแนน STOU-EPT (E-Testing)*. สืบค้นเมื่อวันที่ 29 สิงหาคม 2565 จาก <https://www.stou.ac.th/offices/ore/Adre/STOU-EPT/CEFR.pdf>
- สิริรัตน์ วิภาสศิลป์, ชุศักดิ์ ศิริรุ่งพันธ์, สุนี แข่งเงิน, พิชพันธ์ สาริยันต์, และ ณิชชา มาลัย (2558). *การพัฒนารูปแบบการบริหารการสอบความรู้ภาษาอังกฤษ STOU-EPT ผ่านสื่ออิเล็กทรอนิกส์* (รายงานการวิจัย). นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- สุภมาส อังศุโชติ, สุนิสา จั๋ยม่วงศรี, ปิยนาง สิทธิฤทธิ์, ศจี จิระโร, อนันดา สันธิติวณิชย์, ชุศักดิ์ ศิริรุ่งพันธ์, และ ประรณนา พลอภิชาติ. (2562). *การประเมินระบบการสอบวัดความรู้ภาษาอังกฤษ STOU-EPT ผ่านระบบสอบอิเล็กทรอนิกส์ของมหาวิทยาลัยสุโขทัยธรรมาธิราช* (รายงานการวิจัย). นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- สุริพร อนุศาสนนันท์. (2550). *การเปรียบเทียบคุณภาพของการกำหนดมาตรฐานระหว่างวิธีเอนกอฟ ที่ได้รับการปรับปรุงกับวิธีบูคมาร์ค*. [ปริญญาณิพนธ์ปริญาดุขฉฐิบัณฑิต]. สาขาวิชาการศึกษา และประเมินผลการศึกษา, คณะครุศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย. สืบค้นเมื่อวันที่ 30 สิงหาคม 2565 จาก <http://cuir.car.chula.ac.th/handle/123456789/19722>
- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. Continuum.
- Alderson, J. C. (2007). The CEFR and the need for more research. *The Modern Language Journal*, 91(4), 659–663. <http://www.jstor.org/stable/4626093>
- Al Lawati, Z. A. (2023). Investigating the characteristics of language test specifications and item writer guidelines, and their effect on item development: A mixed-method case study. *Language Testing in Asia*, 13(1), 21. <https://doi.org/10.1186/s40468-023-00233-5>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

- Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational Measurement* (pp. 508-600). Washington, DC: American Council on Education.
- Barman A. (2008). Standard setting in student assessment: Is a defensible method yet to come? *Annals of the Academy of Medicine, Singapore*, 37(11), 957–963.
- Baron, P. A., & Papageorgiou, S. (2016). *Setting language proficiency score requirements for English-as-a-second-language placement decisions in secondary education*. Research Report ETS RR—16–17. Princeton, NJ: Educational Testing Service.
- Bejar, I. I. (2008). Standard setting: What is it? Why is it important? *R&D Connections*, 7, 1–6.
- Berk, R. A. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research*, 56(1), 137–172.
- Blank, R. K. (2002). Using surveys of enacted curriculum to advance evaluation of instruction in relation to standards. *Peabody Journal of Education*, 77(4), 86–121. https://doi.org/10.1207/S15327930PJE7704_5
- British Council & EAQUALS. (2015). *British Council – EAQUALS core inventory for general English* (2nd ed.). <https://www.teachingenglish.org.uk/sites/teacheng/files/pub-british-council-eaquals-core-inventoryv2.pdf>
- Cephe, P. T., & Toprak, T. E. (2014). The Common European Framework of Reference for Languages: Insights for language testing. *Journal of Language and Linguistic Studies*, 10(1), 79–88.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Sage.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge, Cambridge University Press.
- Council of Europe. (2018). CEFR companion volume with new descriptors. Council of Europe.
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Strasbourg: Companion volume, Council of Europe Publishing. Retrieved August 26, 2022, from www.coe.int/lang-cefr.
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press.

- DeMars, C. (2010). *Item response theory*. Oxford University Press.
- Dimitrov, D. M. (2003). The relationship between item response theory and classical test theory: A comparison of the two approaches. *Educational and Psychological Measurement*, 63(6), 1009-1022. <https://doi.org/10.1177/0013164403258032>
- EAQUALS. (2008). EAQUALS bank of descriptors – As checklists. https://www.eaquals.org/wp-content/uploads/EAQUALS_Bank_as_checklists.pdf
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Psychology Press
- Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Brooks/Cole Pub. Co.
- Figueras, N. (2012). The impact of the CEFR. *ELT Journal*, 66(4), 477–485.
- Forte, E. (2017). *Evaluating alignment in large-scale standards-based assessment systems*. Council of Chief State School Officers.
- Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education: Principles, Policy & Practice*, 14(1), 75–97. <https://doi.org/10.1080/09695940701272880>
- Green, A. (2012). *Language functions revisited: Theoretical and empirical bases for language construct definition across the ability range*. Cambridge University Press.
- Gulliksen, H. (1950). *Theory of mental tests*. Wiley.
- Hambleton, R. K. (1978). On the use of cut-scores with criterion-referenced tests in instructional settings. *Journal of Educational Measurement*, 15, 277–290.
- Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 433–470). Praeger.
- Hambleton, R. K., & Plake, B. S. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education*, 8(1), 41–55.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage Publications.

- Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91(4), 663–667. https://doi.org/10.1111/j.1540-4781.2007.00627_5.x
- Impara, J. C., & Plake, B. S. (1997). Standard setting: An alternative approach. *Journal of Educational Measurement*, 34, 353–366.
- Kampa, N., Wagner, H., & Köller, O. (2019). The standard setting process: Validating interpretations of stakeholders. *Large-scale Assessments in Education*, 7(3). <https://doi.org/10.1186/s40536-019-0071-8>
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). American Council on Education & Praeger.
- 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>
- Lee, M. K., Abou Chaar, M. K., Blackmon, S. H., & Yost, K. J. (2024). PRO-Angoff method for remote standard setting: Establishing clinical thresholds for the Upper Digestive Disease Tool. *Journal of Patient-Reported Outcomes*, 8(30). <https://doi.org/10.1186/s41687-024-00707-x>
- Lipnevich, A. A., Khorramdel, L., & Smith, J. K. (2023). Assessment, evaluation, and accountability: A brief introduction. In R. J. Tierney, F. Rizvi, & K. Ercikan (Eds.), *International Encyclopedia of Education* (4th ed., pp. 192–201). Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.09004-7>
- Little, D. (2006). The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. *Language Teaching*, 39(3), 167–190.
- Millman, J. (1973). Passing scores and test lengths for domain-referenced measures. *Review of Educational Research*, 43(2), 205–216. <https://doi.org/10.2307/1169937>
- Neumann, H., Padden, N., & McDonough, K. (2019). Beyond English language proficiency scores: Understanding the academic performance of international undergraduate students during the first year of study. *Higher Education Research & Development*, 38(2), 324–338. <https://doi.org/10.1080/07294360.2018.1522621>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

- Oliver, R., Vanderford, S., & Grote, E. (2012). Evidence of English language proficiency and academic achievement of non-English-speaking background students. *Higher Education Research & Development*, 31(4), 541–555.
<https://doi.org/10.1080/07294360.2011.653958>
- Papageorgiou, S., Tannenbaum, R. J., Bridgeman, B., & Cho, Y. (2015). *The association between TOEFL iBT test scores and academic performance: A meta-analysis. ETS Research Report Series*, 2015(1), 1–46. <https://doi.org/10.1002/ets2.12045>
- Papageorgiou, S., Wu, S., Hsieh, Ching-Ni., Tannenbaum, R. J., & Cheng, M. (2019). *Mapping the TOEFL iBT® test scores to China's standards of English language ability: Implications for score interpretation and use*. ETS Research Report.
- Plake, B. S., & Cizek, G. J. (2012). The modified Angoff, extended Angoff, and Yes/No standard setting methods. In G.J. Cizek (Ed.), *Setting performance standards. Foundations, methods, and innovations* (pp. 181–253). New York: Routledge.
- Polikoff, M.S., Gasparian, H., Korn, S., Gamboa, M., Porter, A.C., Smith, T.M., & Garet, M.S. (2019). Flexibly using the surveys of enacted curriculum to study alignment. *Educational Measurement: Issues and Practice*, 39(2), 38–47. <https://doi.org/10.1111/EMIP.12292>
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>
- Roach, A. T., & Elliott, S. N. (2004). *Alignment analysis and standard-setting procedures for alternate assessments* (WCER Working Paper No. 2004-1). Wisconsin Center for Education Research, University of Wisconsin–Madison.
- Roach, A. T., Elliott, S. N., & Webb, N. L. (2005). Alignment of an alternate assessment with state academic standards: Evidence for the content validity of the Wisconsin alternate assessment. *The Journal of Special Education*, 38(4), 218–231.
- Roach, A. T., Wixson, C., and Talapatra, D. (2010). Aligning an early childhood assessment to state kindergarten content standards: Application of a nationally recognized alignment framework. *Educational Measurement: Issues and Practice*, 29(1), 25–37. <https://doi.org/10.1111/j.1745-3992.2009.00167.x>
- Rothman, R., Slattery, J. B., Vranek, J. L., & Resnick, L. B. (2002). *Benchmarking and alignment of standards and testing*. University of Pittsburgh.

- Rovinelli, R. J., & Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Tijdschrift voor Onderwijsresearch*, 2, 49–60.
- Ruegg, R., Hoang, H., & Petersen, N. (2024). The influence of English language proficiency test scores on the academic success of ESL undergraduate students. *Educational Research and Evaluation*, 29(3-4), 111–129.
<https://doi.org/10.1080/13803611.2024.2314533>
- Tangsakul, S., & Poonpon, K. (2024). Aligning academic reading tests to the Common European Framework of Reference for Languages (CEFR). *rEFLections*, 31(2), 614–638. <https://doi.org/10.61508/refl.v31i2.275057>
- Tannenbaum, R. J. & Baron, P. A. (2011). *Mapping TOEIC® ITP scores onto the Common European Framework of Reference*. Research Memorandum ETS RM—11–33. Princeton, NJ: Educational Testing Service.
- Traynor, A., Webb, N. L., Christopherson, S., & Sato, E. (2020). *Using the results of content alignment analyses to inform ongoing item-level improvements to an assessment program: A guide for state departments of education and for assessment vendors*. Wisconsin Center for Education Products and Services.
- van der Linden, W. J. (2016). *Handbook of item response theory: Volume one: Models*. CRC Press.
- Wainer, H., Bradlow, E. T., & Wang, X. (2007). *Testlet response theory and its applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511490002>
- Wang, L., Chen, X., & Liu, Y. (2022). Setting performance standards for language proficiency tests: A comparative study of methods and implications. *Language Testing*, 39(1), 25–44. <https://doi.org/10.1177/02655322211012345>
- Wang, X., & Wilson, M. (2005). Item response theory with testlets: An overview of the current state of research and future directions. *Educational Measurement: Issues and Practice*, 24(4), 30-40. <https://doi.org/10.1111/j.1745-3992.2005.00021.x>
- Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education*. Council of Chief State School Officers and National Institute for Science Education Research Monograph No. 6. Madison, WI: University of Wisconsin.

- Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied measurement in education*, 20(1), 7-25.
- Yen, W. M. (1983). *Tau equivalence and equipercentile equating*. *Psychometrika*, 48(3), 353-369. <https://doi.org/10.1007/BF02296580>
- Yim, M. K., & Shin, S. (2020). Using the Angoff method to set a standard on mock exams for the Korean Nursing Licensing Examination. *Journal of Educational Evaluation for Health Professions*, 17(14). <https://doi.org/10.3352/jeehp.2020.17.14>

