

Assessment in Health Professions (Semester Year 2025/26)

View Online



[1]

Advancing medical professionalism: 2018.
<https://www.healthcarevalues.ox.ac.uk/files/ampsummarypdf>.

[2]

Anderson, D. and Ackerman-Anderson, L.S. 2010. Beyond change management: how to achieve breakthrough results through conscious change leadership. Pfeiffer.

[3]

Anderson, L.W. and Bloom, B.S. 2001. A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. Longman.

[4]

Archer, J.C. 2005. Use of SPRAT for peer review of paediatricians in training. *BMJ*. 330, 7502 (May 2005), 1251-1253. <https://doi.org/10.1136/bmj.38447.610451.8F>.

[5]

Arnold, L. et al. 2007. Can There Be a Single System for Peer Assessment of Professionalism among Medical Students? A Multi-Institutional Study. *Academic Medicine*. 82, 6 (June 2007), 578-586. <https://doi.org/10.1097/ACM.0b013e3180555d4e>.

[6]

Arnold, L. et al. 2005. Medical students' views on peer assessment of professionalism.

Journal of General Internal Medicine. 20, 9 (Sept. 2005), 819–824.
<https://doi.org/10.1111/j.1525-1497.2005.0162.x>.

[7]

Ashraf, H. et al. 2018. Eye-tracking technology in medical education: A systematic review. *Medical Teacher*. 40, 1 (Jan. 2018), 62–69.
<https://doi.org/10.1080/0142159X.2017.1391373>.

[8]

Association for the Study of Medical Education 2019. *Understanding medical education: evidence, theory, and practice*. Wiley-Blackwell.

[9]

Auewarakul, C. et al. 2005. Item Analysis to Improve Reliability for an Internal Medicine Undergraduate OSCE. *Advances in Health Sciences Education*. 10, 2 (June 2005), 105–113.
<https://doi.org/10.1007/s10459-005-2315-3>.

[10]

Aylward, M. et al. 2014. An Entrustable Professional Activity (EPA) for Handoffs as a Model for EPA Assessment Development. *Academic Medicine*. 89, 10 (Oct. 2014), 1335–1340.
<https://doi.org/10.1097/ACM.0000000000000317>.

[11]

Barbara S. Plake, James C. Impara and Patrick M. Irwin 2000. Consistency of Angoff-Based Predictions of Item Performance: Evidence of Technical Quality of Results from the Angoff Standard Setting Method. *Journal of Educational Measurement*. 37, 4 (2000).

[12]

Ben-David, M.F. 2000. AMEE Guide No. 18: Standard setting in student assessment. *Medical Teacher*. 22, 2 (Jan. 2000), 120–130. <https://doi.org/10.1080/01421590078526>.

[13]

Biggs, J. 1996. Enhancing Teaching through Constructive Alignment. *Higher Education*. 32, 3 (1996), 347–364.

[14]

Bing-You, R.G. 2009. Why Medical Educators May Be Failing at Feedback. *JAMA*. 302, 12 (Sept. 2009). <https://doi.org/10.1001/jama.2009.1393>.

[15]

Black, P. and Wiliam, D. 1998. Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*. 5, 1 (Mar. 1998), 7–74. <https://doi.org/10.1080/0969595980050102>.

[16]

Boursicot, K.A.M. et al. 2018. Structured Assessments of Clinical Competence. *Understanding Medical Education*. T. Swanwick et al., eds. John Wiley & Sons, Ltd. 335–345.

[17]

Brookfield, S. 1987. *Developing critical thinkers: challenging adults to explore alternative ways of thinking and acting*. Open University Press.

[18]

Brown, C. et al. 2015. Money makes the (medical assessment) world go round: The cost of components of a summative final year Objective Structured Clinical Examination (OSCE). *Medical Teacher*. 37, 7 (July 2015), 653–659. <https://doi.org/10.3109/0142159X.2015.1033389>.

[19]

Buckley, S. et al. 2009. The educational effects of portfolios on undergraduate student learning: A Best Evidence Medical Education (BEME) systematic review. BEME Guide No. 11. *Medical Teacher*. 31, 4 (Jan. 2009), 282–298. <https://doi.org/10.1080/01421590902889897>.

[20]

Burls, A. 2009. What is critical appraisal?

[21]

Carraccio, Carol;Wolfsthal, Susan D.;Englander, Robert;Ferentz, Kevin;Martin, Christine
Shifting Paradigms: From Flexner to Competencies. *Academic Medicine*. 77, 5.

[22]

Case, S.M. and Swanson, D.B. 1993. Extended-matching items: A practical alternative to free-response questions. *Teaching and Learning in Medicine*. 5, 2 (Jan. 1993), 107-115.
<https://doi.org/10.1080/10401339309539601>.

[23]

CASP Critical Appraisal Skills Programme Oxford UK: <http://www.casp-uk.net/>.

[24]

ten Cate, O. 2013. Nuts and Bolts of Entrustable Professional Activities. *Journal of Graduate Medical Education*. 5, 1 (Mar. 2013), 157-158. <https://doi.org/10.4300/JGME-D-12-00380.1>.

[25]

ten Cate, O. 2013. Nuts and Bolts of Entrustable Professional Activities. *Journal of Graduate Medical Education*. 5, 1 (Mar. 2013), 157-158. <https://doi.org/10.4300/JGME-D-12-00380.1>.

[26]

ten Cate, O. and Young, J.Q. 2012. The patient handover as an entrustable professional activity: adding meaning in teaching and practice. *BMJ Quality & Safety*. 21, Suppl 1 (Dec. 2012), i9-i12. <https://doi.org/10.1136/bmjqs-2012-001213>.

[27]

Charlin, B. et al. 2000. The Script Concordance Test: A Tool to Assess the Reflective Clinician. *Teaching and Learning in Medicine*. 12, 4 (Oct. 2000), 189–195. https://doi.org/10.1207/S15328015TLM1204_5.

[28]

Cleland, J.A. et al. 2008. Is it me or is it them? Factors that influence the passing of underperforming students. *Medical Education*. 42, 8 (Aug. 2008), 800–809. <https://doi.org/10.1111/j.1365-2923.2008.03113.x>.

[29]

Cobb, K.A. et al. 2013. The educational impact of assessment: A comparison of DOPS and MCQs. *Medical Teacher*. 35, 11 (Nov. 2013), e1598–e1607. <https://doi.org/10.3109/0142159X.2013.803061>.

[30]

Cochrane | Trusted evidence. Informed decisions. Better health.: <http://www.cochrane.org/>.

[31]

Cohen, L. et al. 2018. *Research methods in education*. Routledge.

[32]

Cohen-Schotanus, J. and van der Vleuten, C.P.M. 2010. A standard setting method with the best performing students as point of reference: Practical and affordable. *Medical Teacher*. 32, 2 (Jan. 2010), 154–160. <https://doi.org/10.3109/01421590903196979>.

[33]

Cotton, D.R.E. et al. 2024. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*. 61, 2 (Mar. 2024), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>.

[34]

Courteille, O. et al. 2008. The use of a virtual patient case in an OSCE-based exam – A pilot study. *Medical Teacher*. 30, 3 (Jan. 2008), e66–e76.
<https://doi.org/10.1080/01421590801910216>.

[35]

Cruess, R. The Professionalism Mini-Evaluation Exercise: A Preliminary Investigation. *Academic Medicine*. 81, 10.

[36]

Cruess, R.L. et al. 2016. Amending Miller's Pyramid to Include Professional Identity Formation. *Academic Medicine*. 91, 2 (Feb. 2016), 180–185.
<https://doi.org/10.1097/ACM.0000000000000913>.

[37]

Da Silva, A. and Dennick, R. 2010. Corpus analysis of problem-based learning transcripts: an exploratory study. *Medical Education*. 44, 3 (2010), 280–288.
<https://doi.org/10.1111/j.1365-2923.2009.03575.x>.

[38]

De Champlain, A.F. 2019. Standard Setting Methods in Medical Education. *Understanding Medical Education*. T. Swanwick, ed. Wiley-Blackwell. 347–359.

[39]

De Wever, B. et al. 2011. Assessing collaboration in a wiki: The reliability of university students' peer assessment. *The Internet and Higher Education*. 14, 4 (Sept. 2011), 201–206. <https://doi.org/10.1016/j.iheduc.2011.07.003>.

[40]

Denison, A. et al. 2016. Tablet versus paper marking in assessment: feedback matters. *Perspectives on Medical Education*. 5, 2 (Apr. 2016), 108–113.
<https://doi.org/10.1007/s40037-016-0262-8>.

[41]

Development of generic professional capabilities:
<http://www.gmc-uk.org/education/23581.asp>.

[42]

Devine, O.P. et al. 2015. Assessment at UK medical schools varies substantially in volume, type and intensity and correlates with postgraduate attainment. *BMC Medical Education*. 15, 1 (Dec. 2015). <https://doi.org/10.1186/s12909-015-0428-9>.

[43]

Ditchfield, C. 2007. How do learners make sense of the formative assessment opportunities available to inform their learning in a PBL course. (2007).

[44]

Dory, V. et al. 2012. How to construct and implement script concordance tests: insights from a systematic review. *Medical Education*. 46, 6 (June 2012), 552–563.
<https://doi.org/10.1111/j.1365-2923.2011.04211.x>.

[45]

Downing, S.M. 2003. Guessing on selected-response examinations. *Medical Education*. 37, 8 (Aug. 2003), 670–671. <https://doi.org/10.1046/j.1365-2923.2003.01585.x>.

[46]

Downing, S.M. et al. 2006. RESEARCH METHODOLOGY: Procedures for Establishing Defensible Absolute Passing Scores on Performance Examinations in Health Professions Education. *Teaching and Learning in Medicine*. 18, 1 (Jan. 2006), 50–57.
https://doi.org/10.1207/s15328015t1m1801_11.

[47]

Driessen, E. and van Tartwijk, J. 2013. Portfolios in personal and professional development. *Understanding Medical Education*. T. Swanwick, ed. Wiley-Blackwell. 255–262.

[48]

Driessen, E.W. et al. 2007. Web- or paper-based portfolios: is there a difference? *Medical Education*. 41, 11 (Nov. 2007), 1067–1073.
<https://doi.org/10.1111/j.1365-2923.2007.02859.x>.

[49]

Duers, L.E. and Brown, N. 2009. An exploration of student nurses' experiences of formative assessment. *Nurse Education Today*. 29, 6 (Aug. 2009), 654–659.
<https://doi.org/10.1016/j.nedt.2009.02.007>.

[50]

Durning, S.J. et al. 2002. Assessing the Reliability and Validity of the Mini-Clinical Evaluation Exercise for Internal Medicine Residency Training. *Academic Medicine*. 77, 9 (2002).

[51]

Elnicki, D.M. et al. 1998. Oral versus written feedback in medical clinic. *Journal of General Internal Medicine*. 13, 3 (Mar. 1998), 155–158.
<https://doi.org/10.1046/j.1525-1497.1998.00049.x>.

[52]

Elstein, A.S. et al. *Medical Problem Solving: An Analysis of Clinical Reasoning*.

[53]

Epstein, R.M. 2007. Assessment in Medical Education. *New England Journal of Medicine*. 356, 4 (Jan. 2007), 387–396. <https://doi.org/10.1056/NEJMra054784>.

[54]

Epstein, R.M. 2002. Defining and Assessing Professional Competence. *JAMA*. 287, 2 (Jan. 2002). <https://doi.org/10.1001/jama.287.2.226>.

[55]

Ertmer, P.A. et al. 2007. Using Peer Feedback to Enhance the Quality of Student Online Postings: An Exploratory Study. *Journal of Computer-Mediated Communication*. 12, 2 (Jan. 2007), 412–433. <https://doi.org/10.1111/j.1083-6101.2007.00331.x>.

[56]

Eva, K.W. et al. 2004. An admissions OSCE: the multiple mini-interview. *Medical Education*. 38, 3 (Mar. 2004), 314–326. <https://doi.org/10.1046/j.1365-2923.2004.01776.x>.

[57]

Farmer, E.A. and Page, G. 2005. A practical guide to assessing clinical decision-making skills using the key features approach. *Medical Education*. 39, 12 (Dec. 2005), 1188–1194. <https://doi.org/10.1111/j.1365-2929.2005.02339.x>.

[58]

FENDERSON, B. 1997. The virtues of extended matching and uncued tests as alternatives to multiple choice questions. *Human Pathology*. 28, 5 (May 1997), 526–532. [https://doi.org/10.1016/S0046-8177\(97\)90073-3](https://doi.org/10.1016/S0046-8177(97)90073-3).

[59]

Ferrell, G. 2013. Supporting assessment and feedback practice with technology: from tinkering to transformation.

[60]

Finn, G. et al. 2009. Peer estimation of lack of professionalism correlates with low Conscientiousness Index scores. *Medical Education*. 43, 10 (Oct. 2009), 960–967. <https://doi.org/10.1111/j.1365-2923.2009.03453.x>.

[61]

Fournier, J. et al. 2008. Script Concordance Tests: Guidelines for Construction. *BMC Medical Informatics and Decision Making*. 8, (2008). <https://doi.org/10.1186/1472-6947-8-18>.

[62]

Fowell, S.L. et al. 2008. Estimating the Minimum Number of Judges Required for Test-centred Standard Setting on Written Assessments. Do Discussion and Iteration have an Influence? *Advances in Health Sciences Education*. 13, 1 (Mar. 2008), 11-24. <https://doi.org/10.1007/s10459-006-9027-1>.

[63]

Frost, J. et al. 2012. Expanding assessment methods and moments in history. *Assessment & Evaluation in Higher Education*. 37, 3 (May 2012), 293-304. <https://doi.org/10.1080/02602938.2010.531247>.

[64]

Garrison, C. and Ehringhaus, M. 2007. *Formative and Summative Assessments in the Classroom*. (2007).

[65]

Garrison, D.R. 1991. Critical thinking and adult education: a conceptual model for developing critical thinking in adult learners. *International Journal of Lifelong Education*. 10, 4 (Oct. 1991), 287-303. <https://doi.org/10.1080/0260137910100403>.

[66]

Gaufberg, E. and Fitzpatrick, A. 2008. The favour: a professional boundaries OSCE station. *Medical Education*. 42, 5 (May 2008), 529-530. <https://doi.org/10.1111/j.1365-2923.2008.03067.x>.

[67]

Ginsburg, S. et al. 2004. Basing the evaluation of professionalism on observable behaviours: a cautionary tale. *Academic Medicine*. 79, 10 (2004), S1-S4.

[68]

Ginsburg, S. Context, Conflict, and Resolution: A New Conceptual Framework for Evaluating Professionalism. *Academic Medicine*. 75, 10.

[69]

Ginsburg, S. et al. 2009. From behaviours to attributions: further concerns regarding the evaluation of professionalism. *Medical Education*. 43, 5 (May 2009), 414–425. <https://doi.org/10.1111/j.1365-2923.2009.03335.x>.

[70]

Ginsburg, S. et al. 2016. Hedging to save face: a linguistic analysis of written comments on in-training evaluation reports. *Advances in Health Sciences Education*. 21, 1 (Mar. 2016), 175–188. <https://doi.org/10.1007/s10459-015-9622-0>.

[71]

Goldie, J. 2013. Assessment of professionalism: A consolidation of current thinking. *Medical Teacher*. 35, 2 (Feb. 2013), e952–e956. <https://doi.org/10.3109/0142159X.2012.714888>.

[72]

Gorania, R. 2021. Situational judgement stress. *British Dental Journal*. 231, 8 (Oct. 2021), 426–426. <https://doi.org/10.1038/s41415-021-3577-8>.

[73]

Gould, J. and Day, P. 2013. Hearing you loud and clear: student perspectives of audio feedback in higher education. *Assessment & Evaluation in Higher Education*. 38, 5 (Aug. 2013), 554–566. <https://doi.org/10.1080/02602938.2012.660131>.

[74]

Gravina, E.W. 2017. Competency-Based Education and Its Effect on Nursing Education: A Literature Review. *Teaching and Learning in Nursing*. 12, 2 (Apr. 2017), 117–121. <https://doi.org/10.1016/j.teln.2016.11.004>.

[75]

Guraya, S.Y. 2016. The Desired Concept Maps and Goal Setting for Assessing

Professionalism in Medicine. JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH. (2016). <https://doi.org/10.7860/JCDR/2016/19917.7832>.

[76]

Hagel, C.M. et al. 2016. Queen's University Emergency Medicine Simulation OSCE: an Advance in Competency-Based Assessment. CJEM. 18, 3 (May 2016), 230–233. <https://doi.org/10.1017/cem.2015.34>.

[77]

Haladyna, T.M. et al. 2002. A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. Applied Measurement in Education. 15, 3 (July 2002), 309–333. https://doi.org/10.1207/S15324818AME1503_5.

[78]

Hammick, M. et al. 2010. Conducting a best evidence systematic review. Part 1: From idea to data coding. BEME Guide No. 13. Medical Teacher. 32, 1 (Jan. 2010), 3–15. <https://doi.org/10.3109/01421590903414245>.

[79]

Harden, R.M. et al. 1975. Assessment of clinical competence using objective structured examination. BMJ. 1, 5955 (Feb. 1975), 447–451. <https://doi.org/10.1136/bmj.1.5955.447>.

[80]

Harden, R.M. 2007. Learning outcomes as a tool to assess progression. Medical Teacher. 29, 7 (Jan. 2007), 678–682. <https://doi.org/10.1080/01421590701729955>.

[81]

Harden, R.M. 2015. Misconceptions and the OSCE. Medical Teacher. 37, 7 (July 2015), 608–610. <https://doi.org/10.3109/0142159X.2015.1042443>.

[82]

Harden, R.M. 2016. Revisiting 'Assessment of clinical competence using an objective structured clinical examination (OSCE)'. *Medical Education*. 50, 4 (Apr. 2016), 376–379. <https://doi.org/10.1111/medu.12801>.

[83]

Harden, R.M. et al. 2016. *The definitive guide to the OSCE: the Objective Structured Clinical Examination as a performance assessment*. Elsevier.

[84]

Harrison, C.J. et al. 2015. How we give personalised audio feedback after summative OSCEs. *Medical Teacher*. 37, 4 (Apr. 2015), 323–326. <https://doi.org/10.3109/0142159X.2014.932901>.

[85]

Hauer, K.E. et al. 2013. Developing Entrustable Professional Activities as the Basis for Assessment of Competence in an Internal Medicine Residency: A Feasibility Study. *Journal of General Internal Medicine*. 28, 8 (Aug. 2013), 1110–1114. <https://doi.org/10.1007/s11606-013-2372-x>.

[86]

Hay, D. et al. 2008. Making learning visible: the role of concept mapping in higher education. *Studies in Higher Education*. 33, 3 (June 2008), 295–311. <https://doi.org/10.1080/03075070802049251>.

[87]

Hay, D.B. et al. 2010. Non-traditional learners in higher education: comparison of a traditional MCQ examination with concept mapping to assess learning in a dental radiological science course. *Assessment & Evaluation in Higher Education*. 35, 5 (Aug. 2010), 577–595. <https://doi.org/10.1080/02602931003782525>.

[88]

Hift, R.J. 2014. Should essays and other "open-ended"-type questions retain a place in written summative assessment in clinical medicine? *BMC Medical Education*. 14, 1 (Dec. 2014). <https://doi.org/10.1186/s12909-014-0249-2>.

[89]

Ho, V.W. et al. 2018. Knowledge maps: a tool for online assessment with automated feedback. *Medical Education Online*. 23, 1 (Jan. 2018).
<https://doi.org/10.1080/10872981.2018.1457394>.

[90]

Hodges, B. et al. 1999. OSCE checklists do not capture increasing levels of expertise. 74, 10 (1999).

[91]

Hodges, B. and McIlroy, J.H. 2003. Analytic global OSCE ratings are sensitive to level of training. *Medical Education*. 37, 11 (Nov. 2003), 1012–1016.
<https://doi.org/10.1046/j.1365-2923.2003.01674.x>.

[92]

Hodges, B.D. 2017. *A practical guide for medical teachers*. Elsevier.

[93]

Hodges, B.D. et al. 2011. Assessment of professionalism: Recommendations from the Ottawa 2010 Conference. *Medical Teacher*. 33, 5 (May 2011), 354–363.
<https://doi.org/10.3109/0142159X.2011.577300>.

[94]

Hofstee, W.K.B. 1984. *The Case for Compromise in Educational Selection and Grading. On Educational Testing*. (1984).

[95]

Holmboe, Eric.S. and Durning, S.J. 2024. *Practical Guide to the Assessment of Clinical Competence*.

[96]

Holmboe, E.S. et al. 2003. Construct Validity of the MiniClinical Evaluation Exercise (miniCEX). *Academic Medicine*. 78, 8 (2003).

[97]

Hopwood, J. et al. 2021. Twelve tips for conducting a virtual OSCE. *Medical Teacher*. 43, 6 (June 2021), 633–636. <https://doi.org/10.1080/0142159X.2020.1830961>.

[98]

Horsley, T. et al. 1996. Teaching critical appraisal skills in healthcare settings. *Cochrane Database of Systematic Reviews*. (Sept. 1996). <https://doi.org/10.1002/14651858.CD001270.pub2>.

[99]

Huang, G.C. et al. 2014. Critical Thinking in Health Professions Education: Summary and Consensus Statements of the Millennium Conference 2011. *Teaching and Learning in Medicine*. 26, 1 (Jan. 2014), 95–102. <https://doi.org/10.1080/10401334.2013.857335>.

[100]

Humphrey-Murto, S. et al. 2018. Assessing the Validity of a Multidisciplinary Mini-Clinical Evaluation Exercise. *Teaching and Learning in Medicine*. 30, 2 (Apr. 2018), 152–161. <https://doi.org/10.1080/10401334.2017.1387553>.

[101]

Hurst, Y.K. et al. 2004. The patient assessment questionnaire: A new instrument for evaluating the interpersonal skills of vocational dental practitioners. *British Dental Journal*. 197, 8 (Oct. 2004), 497–500. <https://doi.org/10.1038/sj.bdj.4811750>.

[102]

IC McManus 2014. Implementing statistical equating for MRCP(UK) parts 1 and 2. *BMC Medical Education*. 14, 1 (2014).

[103]

Ingrid Tonni, Cynthia C. Gadbury-Amyot, Marjan Govaerts, Olle ten Cate, Joan Davis, Lily T. Garcia, Richard W. Valachovic 2020. ADEA-ADEE Shaping the Future of Dental Education III. *Journal of Dental Education*. 84, 1 (2020), 97–104. <https://doi.org/10.1002/jdd.12024>.

[104]

J. G. Boyle 2020. Viva la VOSCE? *BMC Medical Education*. 20, 1 (2020).

[105]

Jackson, J.L. et al. 2015. The Quality of Written Feedback by Attendings of Internal Medicine Residents. *Journal of General Internal Medicine*. 30, 7 (July 2015), 973–978. <https://doi.org/10.1007/s11606-015-3237-2>.

[106]

Jenicek, M. 2006. The hard art of soft science: Evidence-Based Medicine, Reasoned Medicine or both? *Journal of Evaluation in Clinical Practice*. 12, 4 (Aug. 2006), 410–419. <https://doi.org/10.1111/j.1365-2753.2006.00718.x>.

[107]

Joanna Briggs Institute QARI: <https://jbi.global/>.

[108]

Jolly, B. 2019. Written Assessment. *Understanding Medical Education*. T. Swanwick, ed. Wiley-Blackwell. 261–261.

[109]

Jolly, B. and Dalton, M.J. 2018. Written Assessment. *Understanding Medical Education*. T. Swanwick et al., eds. John Wiley & Sons, Ltd. 291–317.

[110]

Kakadia, R. et al. 2020. Implementing an online OSCE during the COVID-19 pandemic. *Journal of Dental Education*. (July 2020). <https://doi.org/10.1002/jdd.12323>.

[111]

Karantonis, A. and Sireci, S.G. 2006. The Bookmark Standard-Setting Method: A Literature Review. *Educational Measurement: Issues and Practice*. 25, 1 (Mar. 2006), 4-12. <https://doi.org/10.1111/j.1745-3992.2006.00047.x>.

[112]

Kassab, S.E. et al. 2016. Generalisability theory analyses of concept mapping assessment scores in a problem-based medical curriculum. *Medical Education*. 50, 7 (July 2016), 730-737. <https://doi.org/10.1111/medu.13054>.

[113]

Kee, F. and Bickle, I. 2004. Critical thinking and critical appraisal: the chicken and the egg? *QJM*. 97, 9 (Sept. 2004), 609-614. <https://doi.org/10.1093/qjmed/hch099>.

[114]

Kelly, M. et al. 2012. The Clinical Conscientiousness Index. *Academic Medicine*. 87, 9 (Sept. 2012), 1218-1224. <https://doi.org/10.1097/ACM.0b013e3182628499>.

[115]

Kessel, D. et al. 2012. Workplace based assessments are no more. *BMJ*. (Sept. 2012). <https://doi.org/10.1136/bmj.e6193>.

[116]

Kirkpatrick, D. 1996. Great Ideas Revisited: Revisiting Kirkpatrick's Four-Level Model. *Training and Development*. 50, 1 (1996), 54-59.

[117]

Kogan, J.R. et al. 2003. Feasibility, Reliability, and Validity of the Mini-Clinical Evaluation

Exercise (mCEX) in a Medicine Core Clerkship. *Academic Medicine*. 78, 10 (2003).

[118]

Lane, P. 2005. Recruitment into training for general practice—the winds of change or a breath of fresh air? *BMJ*. 331, 7520 (Oct. 2005), s153–s153.
<https://doi.org/10.1136/bmj.331.7520.s153>.

[119]

Linn, A.M.J. et al. 2013. Standard setting of script concordance tests using an adapted Nedelsky approach. *Medical Teacher*. 35, 4 (Apr. 2013), 314–319.
<https://doi.org/10.3109/0142159X.2012.746446>.

[120]

Lubarsky, S. et al. 2018. Examining the effects of gaming and guessing on script concordance test scores. *Perspectives on Medical Education*. 7, 3 (June 2018), 174–181.
<https://doi.org/10.1007/s40037-018-0435-8>.

[121]

Ma, I.W.Y. et al. 2012. Comparing the use of global rating scale with checklists for the assessment of central venous catheterization skills using simulation. *Advances in Health Sciences Education*. 17, 4 (Oct. 2012), 457–470.
<https://doi.org/10.1007/s10459-011-9322-3>.

[122]

Marshall, S. ed. 2020. *A handbook for teaching and learning in higher education: enhancing academic practice*. Routledge.

[123]

Martinsen, S.S.S. et al. 2021. Examining the educational impact of the mini-CEX: a randomised controlled study. *BMC Medical Education*. 21, 1 (Dec. 2021).
<https://doi.org/10.1186/s12909-021-02670-3>.

[124]

Mayer, R.E. 2012. Cognitive Learning. Encyclopedia of the sciences of learning. Springer.

[125]

McCormack, W.T. et al. 2007. Peer Nomination: A Tool for Identifying Medical Student Exemplars in Clinical Competence and Caring, Evaluated at Three Medical Schools. *Academic Medicine*. 82, 11 (Nov. 2007), 1033–1039.
<https://doi.org/10.1097/01.ACM.0000285345.75528.ee>.

[126]

McCoubrie, P. 2004. Improving the fairness of multiple-choice questions: a literature review. *Medical Teacher*. 26, 8 (Dec. 2004), 709–712.
<https://doi.org/10.1080/01421590400013495>.

[127]

McKinley, D.W. and Norcini, J.J. 2014. How to set standards on performance-based examinations: AMEE Guide No. 85. *Medical Teacher*. 36, 2 (Feb. 2014), 97–110.
<https://doi.org/10.3109/0142159X.2013.853119>.

[128]

McLachlan, J.C. et al. 2009. The Conscientiousness Index: A Novel Tool to Explore Students' Professionalism. *Academic Medicine*. 84, 5 (May 2009), 559–565.
<https://doi.org/10.1097/ACM.0b013e31819fb7ff>.

[129]

Meskell, P. et al. 2015. Back to the future: An online OSCE Management Information System for nursing OSCEs. *Nurse Education Today*. 35, 11 (Nov. 2015), 1091–1096.
<https://doi.org/10.1016/j.nedt.2015.06.010>.

[130]

Miller, G. 1990. The assessment of clinical skills/competence/performance. *Academic Medicine*. 65, 9 (1990).

[131]

Miller, M.D. et al. 2013. Measurement and assessment in teaching. Pearson Education.

[132]

Missimer, C.A. 1995. Good arguments: an introduction to critical thinking. Prentice Hall.

[133]

Mitchell, C. et al. 2011. Workplace-based assessments of junior doctors: do scores predict training difficulties? *Medical Education*. 45, 12 (Dec. 2011), 1190–1198.
<https://doi.org/10.1111/j.1365-2923.2011.04056.x>.

[134]

van Mook, W.N.K.A. et al. 2009. Approaches to professional behaviour assessment: Tools in the professionalism toolbox. *European Journal of Internal Medicine*. 20, 8 (Dec. 2009), e153–e157. <https://doi.org/10.1016/j.ejim.2009.07.012>.

[135]

van Mook, W.N.K.A. et al. 2009. General considerations regarding assessment of professional behaviour. *European Journal of Internal Medicine*. 20, 4 (July 2009), e90–e95. <https://doi.org/10.1016/j.ejim.2008.11.011>.

[136]

Moore, T.J. 2011. Critical thinking and disciplinary thinking: a continuing debate. *Higher Education Research & Development*. 30, 3 (June 2011), 261–274. <https://doi.org/10.1080/07294360.2010.501328>.

[137]

Murphy, D.J. et al. 2009. The reliability of workplace-based assessment in postgraduate medical education and training: a national evaluation in general practice in the United Kingdom. *Advances in Health Sciences Education*. 14, 2 (May 2009), 219–232. <https://doi.org/10.1007/s10459-008-9104-8>.

[138]

NEWBLE, D.I. and SWANSON, D.B. 1988. Psychometric characteristics of the objective structured clinical examination. *Medical Education*. 22, 4 (July 1988), 325–334. <https://doi.org/10.1111/j.1365-2923.1988.tb00761.x>.

[139]

Nicol, D.J. and Macfarlane-Dick, D. 2006. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*. 31, 2 (Apr. 2006), 199–218. <https://doi.org/10.1080/03075070600572090>.

[140]

Noel, G.L. 1992. How Well Do Internal Medicine Faculty Members Evaluate the Clinical Skills of Residents? *Annals of Internal Medicine*. 117, 9 (Nov. 1992). <https://doi.org/10.7326/0003-4819-117-9-757>.

[141]

Norcini, J. et al. 2011. Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference. *Medical Teacher*. 33, 3 (Mar. 2011), 206–214. <https://doi.org/10.3109/0142159X.2011.551559>.

[142]

Norcini, J. and Burch, V. 2007. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical Teacher*. 29, 9–10 (Jan. 2007), 855–871. <https://doi.org/10.1080/01421590701775453>.

[143]

Norcini, J.J. 2003. Peer assessment of competence. *Medical Education*. 37, 6 (June 2003), 539–543. <https://doi.org/10.1046/j.1365-2923.2003.01536.x>.

[144]

Norcini, J.J. 2003. The Mini-CEX: A Method for Assessing Clinical Skills. *Annals of Internal*

Medicine. 138, 6 (Mar. 2003). <https://doi.org/10.7326/0003-4819-138-6-200303180-00012>.

[145]

Norcross, W.A. 1985. The Consultation: An Approach to Learning and Teaching. *JAMA: The Journal of the American Medical Association*. 253, 3 (Jan. 1985). <https://doi.org/10.1001/jama.1985.03350270123038>.

[146]

Olson, B.L. and McDonald, J.L. 2004. Influence of Online Formative Assessment Upon Student Learning in Biomedical Science Courses. *Journal of Dental Education*. 68, 6 (June 2004), 656-659. <https://doi.org/10.1002/j.0022-0337.2004.68.6.tb03783.x>.

[147]

Orsini, C. and Binnie, V.I. 2016. Entrustment decisions in dental education: Is it time to start formalising? *Medical Teacher*. 38, 3 (Mar. 2016), 322-322. <https://doi.org/10.3109/0142159X.2015.1114598>.

[148]

Palmer, E.J. and Devitt, P.G. 2007. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions?: research paper. *BMC Medical Education*. 7, 1 (2007). <https://doi.org/10.1186/1472-6920-7-49>.

[149]

Paniagua, M. and Swygert, K. eds *The Gold Book - constructing written test questions for the basic and clinical sciences*.

[150]

Papadakis, M.A. et al. 2005. Disciplinary Action by Medical Boards and Prior Behavior in Medical School. *New England Journal of Medicine*. 353, 25 (Dec. 2005), 2673-2682. <https://doi.org/10.1056/NEJMsa052596>.

[151]

Patterson, F. et al. 2016. Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. *Medical Teacher*. 38, 1 (Jan. 2016), 3–17. <https://doi.org/10.3109/0142159X.2015.1072619>.

[152]

Paul, R. 1995. *Critical thinking: how to prepare students for a rapidly changing world. foundation for critical thinking.*

[153]

Paul, R. and Elder, L. 2006. *The Miniature Guide to Critical Thinking: Concepts and Tools.*

[154]

Peters, H. et al. 2017. Twelve tips for the implementation of EPAs for assessment and entrustment decisions. *Medical Teacher*. 39, 8 (Aug. 2017), 802–807. <https://doi.org/10.1080/0142159X.2017.1331031>.

[155]

Piedra, N. et al. 2010. Measuring collaboration and creativity skills through rubrics: Experience from UTPL collaborative social networks course. *IEEE EDUCON 2010 Conference (2010)*, 1511–1516.

[156]

Pohl, C.A. et al. 2011. Peer Nominations as Related to Academic Attainment, Empathy, Personality, and Specialty Interest. *Academic Medicine*. 86, 6 (June 2011), 747–751. <https://doi.org/10.1097/ACM.0b013e318217e464>.

[157]

Puryer, J. and O'Sullivan, D. 2015. An introduction to standard setting methods in dentistry. *BDJ*. 219, 7 (Oct. 2015), 355–358. <https://doi.org/10.1038/sj.bdj.2015.755>.

[158]

Quantrill, S.J. and Tun, J.K. 2012. Workplace-based assessment as an educational tool. Guide supplement 31.5 – Viewpoint. *Medical Teacher*. 34, 5 (May 2012), 417–418. <https://doi.org/10.3109/0142159X.2012.668234>.

[159]

Ramani, S. and Krackov, S.K. 2012. Twelve tips for giving feedback effectively in the clinical environment. *Medical Teacher*. 34, 10 (Oct. 2012), 787–791. <https://doi.org/10.3109/0142159X.2012.684916>.

[160]

Ramsey, P.G. 1993. Use of Peer Ratings to Evaluate Physician Performance. *JAMA: The Journal of the American Medical Association*. 269, 13 (Apr. 1993). <https://doi.org/10.1001/jama.1993.03500130069034>.

[161]

Read, E.K. et al. 2015. The Use of Global Rating Scales for OSCEs in Veterinary Medicine. *PLOS ONE*. 10, 3 (Mar. 2015). <https://doi.org/10.1371/journal.pone.0121000>.

[162]

Regehr G1, MacRae H, Reznick RK, Szalay D. 1998. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med*. (1998).

[163]

Rekman, J. et al. 2016. A New Instrument for Assessing Resident Competence in Surgical Clinic: The Ottawa Clinic Assessment Tool. *Journal of Surgical Education*. 73, 4 (July 2016), 575–582. <https://doi.org/10.1016/j.jsurg.2016.02.003>.

[164]

Richstone, L. et al. 2010. Eye Metrics as an Objective Assessment of Surgical Skill. *Annals of Surgery*. 252, 1 (July 2010), 177–182. <https://doi.org/10.1097/SLA.0b013e3181e464fb>.

[165]

Rimmer, A. 2023. Situational judgment test is scrapped under new system for allocating foundation training places. *BMJ*. (June 2023). <https://doi.org/10.1136/bmj.p1269>.

[166]

Ross, M. 2015. Entrustable professional activities. *The Clinical Teacher*. 12, 4 (Aug. 2015), 223–225. <https://doi.org/10.1111/tct.12436>.

[167]

Royal College of Physicians 2005. *Doctors in Society: Medical professionalism in a changing world*.

[168]

Rudolph, J. et al. 2013. We know what they did wrong, but not why : the case for 'frame-based' feedback. *The Clinical Teacher*. 10, 3 (June 2013), 186–189. <https://doi.org/10.1111/j.1743-498X.2012.00636.x>.

[169]

Rushforth, H.E. 2007. Objective structured clinical examination (OSCE): Review of literature and implications for nursing education. *Nurse Education Today*. 27, 5 (July 2007), 481–490. <https://doi.org/10.1016/j.nedt.2006.08.009>.

[170]

Rushton, A. 2005. Formative assessment: a key to deep learning? *Medical Teacher*. 27, 6 (Sept. 2005), 509–513. <https://doi.org/10.1080/01421590500129159>.

[171]

Ryan, A. et al. 2020. Fully online OSCEs: A large cohort case study. *MedEdPublish*. 9, 1 (2020). <https://doi.org/10.15694/mep.2020.000214.1>.

[172]

dos S. Ribeiro, C. et al. 2019. Overcoming challenges for designing and implementing the One Health approach: A systematic review of the literature. *One Health*. 7, (June 2019). <https://doi.org/10.1016/j.onehlt.2019.100085>.

[173]

Sadler, D.R. 1989. Formative assessment and the design of instructional systems. *Instructional Science*. 18, 2 (June 1989), 119–144. <https://doi.org/10.1007/BF00117714>.

[174]

Sadler, D.R. 1998. Formative Assessment: revisiting the territory. *Assessment in Education: Principles, Policy & Practice*. 5, 1 (Mar. 1998), 77–84. <https://doi.org/10.1080/0969595980050104>.

[175]

Sam, A.H. et al. 2018. Very-short-answer questions: reliability, discrimination and acceptability. *Medical Education*. 52, 4 (Apr. 2018), 447–455. <https://doi.org/10.1111/medu.13504>.

[176]

Scally, G. and Donaldson, L.J. 1998. Looking forward: Clinical governance and the drive for quality improvement in the new NHS in England. *BMJ*. 317, 7150 (July 1998), 61–65. <https://doi.org/10.1136/bmj.317.7150.61>.

[177]

Schoonheim-Klein, M. et al. 2009. Who will pass the dental OSCE? Comparison of the Angoff and the borderline regression standard setting methods. *European Journal of Dental Education*. 13, 3 (Aug. 2009), 162–171. <https://doi.org/10.1111/j.1600-0579.2008.00568.x>.

[178]

Schubert, S. et al. 2008. A situational judgement test of professional behaviour: development and validation. *Medical Teacher*. 30, 5 (Jan. 2008), 528–533. <https://doi.org/10.1080/01421590801952994>.

[179]

Schuwirth, L.W. and van der Vleuten, C.P. 2019. How to Design a Useful Test: The Principles of Assessment. *Understanding Medical Education*. T. Swanwick, ed. Wiley-Blackwell. 277–289.

[180]

Schuwirth, L.W.T. and van der Vleuten, C.P.M. 2003. ABC of learning and teaching in medicine: Written assessment. *BMJ*. 326, 7390 (Mar. 2003), 643–645. <https://doi.org/10.1136/bmj.326.7390.643>.

[181]

Schuwirth, L.W.T. and van der Vleuten, C.P.M. 2004. Different written assessment methods: what can be said about their strengths and weaknesses? *Medical Education*. 38, 9 (Sept. 2004), 974–979. <https://doi.org/10.1111/j.1365-2929.2004.01916.x>.

[182]

Schuwirth, L.W.T. and van der Vleuten, C.P.M. 2011. General overview of the theories used in assessment: AMEE Guide No. 57. *Medical Teacher*. 33, 10 (Oct. 2011), 783–797. <https://doi.org/10.3109/0142159X.2011.611022>.

[183]

Schuwirth, L.W.T. and van der Vleuten, C.P.M. 2012. Programmatic assessment and Kane's validity perspective. *Medical Education*. 46, 1 (Jan. 2012), 38–48. <https://doi.org/10.1111/j.1365-2923.2011.04098.x>.

[184]

Sender Liberman, A. et al. 2005. Surgery residents and attending surgeons have different perceptions of feedback. *Medical Teacher*. 27, 5 (Aug. 2005), 470–472. <https://doi.org/10.1080/0142590500129183>.

[185]

Shaw, S. 2005. Research governance: where did it come from, what does it mean? *Journal of the Royal Society of Medicine*. 98, 11 (Nov. 2005), 496–502.

[186]

Siau, K. et al. 2018. Changes in scoring of Direct Observation of Procedural Skills (DOPS) forms and the impact on competence assessment. *Endoscopy*. 50, 08 (Aug. 2018), 770–778. <https://doi.org/10.1055/a-0576-6667>.

[187]

Snell, L.S. and Frank, J.R. 2010. Competencies, the tea bag model, and the end of time. *Medical Teacher*. 32, 8 (Aug. 2010), 629–630. <https://doi.org/10.3109/0142159X.2010.500707>.

[188]

Spielman, A. et al. 2005. Dentistry, Nursing, and Medicine: A Comparison of Core Competencies. *Journal of Dental Education*. 69, 11 (2005), 1257–1271.

[189]

Stern, D.T. 2006. *Measuring medical professionalism*. Oxford University Press.

[190]

Stern, D.T. et al. 2005. The prediction of professional behaviour. *Medical Education*. 39, 1 (Jan. 2005), 75–82. <https://doi.org/10.1111/j.1365-2929.2004.02035.x>.

[191]

Sturpe, D.A. 2010. Objective Structured Clinical Examinations in Doctor of Pharmacy Programs in the United States. *American journal of pharmaceutical education*. 74, 8 (2010). <https://doi.org/10.5688/aj7408148>.

[192]

Suetsugu, N. et al. 2016. Quantitative Analysis of Nursing Observation Employing a

Portable Eye-Tracker. *Open Journal of Nursing*. 06, 01 (2016), 53–61.
<https://doi.org/10.4236/ojn.2016.61006>.

[193]

Sutherland, R.M. et al. 2019. Assessing Diagnostic Reasoning Using a Standardized Case-Based Discussion. *Journal of Medical Education and Curricular Development*. 6, (Jan. 2019). <https://doi.org/10.1177/2382120519849411>.

[194]

Tavakol, M. and Doody, G.A. 2016. A novel psychometric programme for the rapid analysis of OSCE data. *Medical Teacher*. 38, 1 (Jan. 2016), 104–105.
<https://doi.org/10.3109/0142159X.2015.1062085>.

[195]

Taylor, C.A. 2011. Development of a modified Cohen method of standard setting. *Medical Teacher*. 33, 12 (Dec. 2011), e678–e682. <https://doi.org/10.3109/0142159X.2011.611192>.

[196]

Tekian, A. et al. 2020. Entrustment decisions: Implications for curriculum development and assessment. *Medical Teacher*. (Mar. 2020), 1–7.
<https://doi.org/10.1080/0142159X.2020.1733506>.

[197]

Ten Cate, O. 2017. Competency-Based Postgraduate Medical Education: Past, Present and Future. *GMS Journal for Medical Education*. 34, 5 (2017).
<https://doi.org/10.3205/zma001146>.

[198]

The Campbell Collaboration: <http://www.campbellcollaboration.org/>.

[199]

Torsney, K.M. et al. 2015. The Modern Surgeon and Competency Assessment: Are the Workplace-Based Assessments Evidence-Based? *World Journal of Surgery*. 39, 3 (Mar. 2015), 623–633. <https://doi.org/10.1007/s00268-014-2875-6>.

[200]

Van De Ridder, J.M.M. et al. 2008. What is feedback in clinical education? *Medical Education*. 42, 2 (Jan. 2008), 189–197. <https://doi.org/10.1111/j.1365-2923.2007.02973.x>.

[201]

Van Der Vleuten, C.P.M. 1996. The assessment of professional competence: Developments, research and practical implications. *Advances in Health Sciences Education* . 1, 1 (Jan. 1996), 41–67. <https://doi.org/10.1007/BF00596229>.

[202]

Verkerk, M.A. et al. 2007. Reflective professionalism: interpreting CanMEDS' 'professionalism'. *Journal of Medical Ethics*. 33, 11 (Nov. 2007), 663–666. <https://doi.org/10.1136/jme.2006.017954>.

[203]

van der Vleuten, C.P.M. et al. 2012. A model for programmatic assessment fit for purpose. *Medical Teacher*. 34, 3 (Mar. 2012), 205–214. <https://doi.org/10.3109/0142159X.2012.652239>.

[204]

van der Vleuten, C.P.M. and Schuwirth, L.W.T. 2005. Assessing professional competence: from methods to programmes. *Medical Education*. 39, 3 (Mar. 2005), 309–317. <https://doi.org/10.1111/j.1365-2929.2005.02094.x>.

[205]

Voelkel, S. and Mello, L.V. 2014. Audio Feedback – Better Feedback? *Bioscience Education*. 22, 1 (July 2014), 16–30. <https://doi.org/10.11120/beej.2014.00022>.

[206]

Watson, R. et al. 2002. Clinical competence assessment in nursing: a systematic review of the literature. *Journal of Advanced Nursing*. 39, 5 (Sept. 2002), 421–431.
<https://doi.org/10.1046/j.1365-2648.2002.02307.x>.

[207]

Weaver, M.R. 2006. Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education*. 31, 3 (June 2006), 379–394.
<https://doi.org/10.1080/02602930500353061>.

[208]

Wilkinson, T.J. et al. 2009. A Blueprint to Assess Professionalism: Results of a Systematic Review. *Academic Medicine*. 84, 5 (May 2009), 551–558.
<https://doi.org/10.1097/ACM.0b013e31819fbaa2>.

[209]

Williams, D.M. et al. 2016. Peer and near-peer OSCE examiners. *Medical Teacher*. 38, 2 (Feb. 2016), 212–213. <https://doi.org/10.3109/0142159X.2015.1072266>.

[210]

Williams, R.G. et al. 2005. Assuring the reliability of resident performance appraisals: More items or more observations? *Surgery*. 137, 2 (Feb. 2005), 141–147.
<https://doi.org/10.1016/j.surg.2004.06.011>.

[211]

Wood, D.F. 2019. Formative Assessment. *Understanding Medical Education*. T. Swanwick, ed. John Wiley & Sons, Ltd. 317–328.

[212]

Wood, T.J. et al. 2006. Standard Setting in a Small Scale OSCE: A Comparison of the Modified Borderline-Group Method and the Borderline Regression Method. *Advances in Health Sciences Education*. 11, 2 (May 2006), 115–122.
<https://doi.org/10.1007/s10459-005-7853-1>.

[213]

Wood, T.J. and Pugh, D. 2020. Are rating scales really better than checklists for measuring increasing levels of expertise? *Medical Teacher*. 42, 1 (Jan. 2020), 46–51. <https://doi.org/10.1080/0142159X.2019.1652260>.

[214]

Woodhouse, L Comparison of Cohen and Angoff methods of standard setting: is Angoff worth it? *European Board of Medical Assessors Annual Academic Conference: Crossing Boundaries: Assessment in Medical Education*.

[215]

Yardley, S. and Dornan, T. 2012. Kirkpatrick's levels and education 'evidence'. *Medical Education*. 46, 1 (Jan. 2012), 97–106. <https://doi.org/10.1111/j.1365-2923.2011.04076.x>.

[216]

Zijlstra-Shaw, S. et al. 2012. Assessing professionalism within dental education; the need for a definition. *European Journal of Dental Education*. 16, 1 (Feb. 2012), e128–e136. <https://doi.org/10.1111/j.1600-0579.2011.00687.x>.

[217]

2010. *Dimensions of Quality*.