BIBLIOGRAFIE

American Association for the Advancement of Science (2017). *AAAS science assessment,* 1st June. http://assessment.aaas.org/topics

Anderson, R., Pichert, J. and Shirey, L. (1983). Effects of the reader's schema at different points in time. *Journal of Educational Psychology*, 75(2), pp. 271–279.

Ariely, D. (2013). *The honest truth about dishonesty: How we lie to everyone – especially ourselves*. New York: Harper.

Athanases, S. and Achinstein, B. (2003). Focusing new teachers on individual and low performing students: The centrality of formative assessment in the mentor's repertoire of practice. *Teachers College Record*, 105(8), pp. 1486–1520.

Aubrey-Hopkins, J. and James, C. (2002). Improving practice in subject departments: The experience of secondary school subject leaders in Wales. *School Leadership & Management*, 22(3), pp. 305–320.

Bahník, Š. and Vranka, M. (2017). Growth mindset is not associated with scholastic aptitude in a large sample of university applicants. *Personality and Individual Differences,* 117, pp. 139–143.

Bailin, S., Case, R., Coombs, J. and Daniels, L. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies,* 31(3), pp. 285–302.

Baird, J., Andrich, D., Hopfenbeck, T. and Stobart, G. (2017). Assessment and learning: Fields apart? *Assessment in Education: Principles, Policy & Practice,* 24(3), pp. 317–350.

Ball, D. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal*, 93(4), pp. 373–397.

Ball, D., Thames, M. and Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education,* 59(5), pp. 389–407.

Bambrick-Santoyo, P. (2016). *Get better faster: A 90-day pan for coaching new teachers*. San Francisco, CA: John Wiley and Sons.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist,* 37(2), pp. 122 – 147.

Bandura, A. and Schunk, D. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, 41(3), pp. 586–598.

Bangert-Drowns, R., Kulik, C., Kulik, J. and Morgan, M. (1991). The instructional effect of feedback in test like events. *Review of Educational Research*, 61(2), pp. 213–238.

Banks, B. (1991). *The KMP way to learn maths: A history of the early development of the Kent Mathematics Project*. Maidstone, UK: Bertram Banks.

Bart, W. M., Post, T., Behr, M. J. and Lesh, R. (1994). A diagnostic analysis of a proportional reasoning test item: An introduction to the properties of a semi-dense item. *Focus on Learning Problems in Mathematics,* 16(3), pp. 1–11.

Behavioural Insights Team (2015). *The Behavioural Insights Team: Update report 2013–2015*. Behavioural Insights Team (2017). *The Behavioural Insights Team: Update report 2016–17*.

Bennett, R. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice,* 18(1), pp. 5–25.

Berger, R. (2003). *An ethic of excellence: Building a culture of craftsmanship with students*. Portsmouth, NH: Heinemann.

Black, P. and Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice,* 5(1), pp. 7–74.

Black, P. and Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. London: GL Assessment.

Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In Tyler, R. W. (ed.), *Educational evaluation: New roles, new means: The 68th yearbook of the national society for the study of education (part II)*, 68(2). Chicago, IL: University of Chicago Press, pp. 26–50.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), pp. 4–16.

Brinkman, S., Johnson, S., Codde, J., Hart, M., Straton, J., Mittinty, M. and Silburn, S. (2016). Efficacy of infant simulator programmes to prevent teenage pregnancy: A school-based cluster randomised controlled trial in Western Australia. *The Lancet*, 388(10057), pp. 2264–2271.

Brown, M., Rutherford, D. and Boyle, B. (2000). Leadership for school improvement: The role of the head of department in UK secondary schools. *School Effectiveness and School Improvement,* 11(2), pp. 237–258.

Brown, P. , Roediger, H. and McDaniel, M. (2014). *Make it stick*. Cambridge, MA: The Belknap Press of Harvard University Press.

Bryk, A. and Schneider, B. (2002). *Trust in schools*. New York: Russell Sage Foundation.

Bubb, S. and Earley, P. (2004). Why is managing change not easy? In *Managing teacher workload: Workload and wellbeing*. London: PCP/Sage.

Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and egoinvolving evaluation on interest and performance. *British Journal of Educational Psychology*, 58(1), pp. 1–14.

Cambridge Assessment (2017). *A Cambridge approach to improving education: Using international insights to manage complexity*.

Casselman, B. and Atwood, C. (2017). Improving general chemistry course performance through online homework-based metacognitive training. *Journal of Chemical Education,* 94(12), pp. 1811–1821. DOI:10.1021/acs.jchemed.7b00298.

Chapman, C. (2004). Leadership for improvement in urban and challenging contexts. *London Review of Education,* 2(2), pp. 95–108.

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In Vosniadou, S. (ed.), *Handbook of research on conceptual change*. Hillsdale, NJ: Erlbaum, pp. 61–82.

Chi, M., Glaser, R. and Rees, E. (1982). Expertise in problem solving. In Sternberg, R. (ed.), *Advances in the psychology of human intelligence*. Hillsdale, NJ: Erlbaum, pp. 7–75.

Christodoulou, D. (2017). *Making good progress: The future of Assessment for Learning*. Oxford: Oxford University Press.

Coe, R. (2013). *Improving education: A triumph of hope over experience*. Centre for Evaluation and Monitoring.

Coe, R., Aloisi, C., Higgins, S. and Elliot Major, L. (2014). *What makes great teaching? Review of the underpinning research*. Sutton Trust.

Coffey, J., Hammer, D., Levin, D. and Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching,* 48(10), pp. 1109–1136.

Conway, R. (2017). Reflecting on ... student responses to feedback. *JMS Reflect* [blog], 18th June. https://jmsreflect.blog/2017/06/18/reflecting-on-student-responses-to-feedback/

Cordingley, P. , Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., Saunders, L. and Coe, R. (2015). *Developing great teaching: Lessons from the international reviews into effective professional development*. London: Teacher Development Trust.

Covey, S. (2003). *The seven habits of highly effective people*. Carlsbad, CA: Hay House.

Cowan, N. (1999). An embedded-processes model of working memory. In Miyake, A. and Shah, P. (eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press, pp. 62–101.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), pp. 438–481.

Crouch, C. H., Watkins, J., Fagen, A. P. and Mazur, E. (2007). Peer instruction: Engaging students one-onone, all at once. *Research-Based Reform of University Physics,* 1(1), pp. 40–95.

Csikszentmihalyi, M. (2002). *Flow: The classic work on how to achieve happiness*. London: Rider.

Darling-Hammond, L., Hammerness, K., Grossman, P. , Rust, F. and Shulman, L. (2005). The design of teacher education programs. In Darling-Hammond, L. and Bransford, J. (eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco: Wiley, Ch. 10.

Darling-Hammond, L., Hyler, M. E. and Gardner, M. (2017). *Effective teacher professional development*. Palo Alto, CA: Learning Policy Institute.

Deans for Impact (2015). *The science of learning*. Austin, TX: Deans for Impact.

Deans for Impact (2016). *Practice with purpose: The emerging science of teacher expertise*. Austin, TX: Deans for Impact.

De Bruyckere, P. , Kirschner, P. and Hulshof, C. (2015). *Urban myths about learning and education*. London: Academic Press.

Department for Education (2015). *Final report of the Commission on Assessment without levels*.

Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Education Researcher,* 38(3), pp. 181–199.

Didau, D. (2014). Why AfL might be wrong, and what to do about. *The learning spy* [blog], 12th March. www.learningspy.co.uk/myths/afl -might-wrong/

Didau, D. and Rose, N. (2016). *What every teacher needs to know about … psychology*. Woodbridge: John Catt.

D'Mello, S., Lehman, B., Pekrun, R. and Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction,* 29, pp. 153–170.

Duit, R. (2009). Bibliography: Students' and teachers' conceptions and science education [online]. http://archiv.ipn.uni-kiel.de/stcse/

Elliott, V., Baird, J., Hopfenbeck, T., Ingram, J., Richardson, J., Coleman, R. Thompson, I., Usher, N. and Zantout, M. (2016). *A marked improvement? A review of the evidence on written marking*. London: Education Endowment Fund.

Ericsson, A. and Pool, R. (2016). *Peak: Secrets from the new science of expertise*. London: Bodley Head.

Facer, J. (2016). Giving feedback the Michaela way. *Reading all the books* [blog], 19th March. https://readingallthebooks.com/2016/03/19/giving-feedback-the-michaela-way/

Feynman, R. (1974). *Cargo cult science* [Commencement address]. Caltech. http://calteches.library.caltech.edu/51/2/CargoCult.htm

Fletcher-Wood, H. (2016). *Ticked off: Checklists for students, teachers and school leaders*. Carmarthen: Crown House.

Fryer, R. (2017). *Management and student achievement: Evidence from a randomized field experiment*. NBER Working Paper No. 23437.

Fuchs, L. S. and Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children,* 53(3), pp. 199–208.

Fuller, F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 6(2), pp. 207–226.

Gawande, A. (2010). *The checklist manifesto: How to get things right*. London: Profile.

Gentner, D. (1976). The structure and recall of narrative prose. *Journal of Verbal Learning and Verbal Behavior,* 15(4), pp. 411–418.

Gibbons, R. F. (1975). An account of the secondary mathematics individualized learning experiment. *Mathematics in School,* 4(6), pp. 14–16.

Gibson, S., Oliver, L. and Dennison, M. (2015). *Workload challenge: Analysis of teacher consultation responses*. Department for Education.

Gick, M. and Holyoak, K. (1980). Analogical problem solving. *Cognitive Psychology,* 12(3), pp. 306–355.

Gierl, M., Bulut, O., Guo, Q. and Zhang, X. (2017). Developing, analyzing, and using distractors for multiple-choice tests in education: A comprehensive review. *Review of Educational Research,* 87(6), pp. 1082–1116.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. Abingdon: Routledge.

Goodrich, J. (2017). Is formative assessment fatally flawed? Confusing learning & performance. *Improving teaching* [blog comment], 26th March. https://improvingteaching.co.uk/2017/03/26/formative-assessment-flawed-confusing-learning-performance/

Gray, E. and Tall, D. (1994). Duality, ambiguity and flexibility: A proceptual view of simple arithmetic. *Journal for Research in Mathematics Education,* 26(2), pp. 115–141.

Grissom, J., Loeb, S. and Master, B. (2013). Effective instructional time use for school leaders: Longitudinal evidence from observations of principals. *Educational Researcher,* 42(8), pp. 433–444.

Grossman, P. , Wineburg, S. and Woolworth, S. (2001). Toward a theory of teacher community. *The Teachers College Record*, 103, pp. 942–1012.

Guskey, T. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice,* 8(3/4), pp. 381–391.

Haladyna, T., Downing, S. and Rodriguez, M. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education,* 15(3), pp. 309–333.

Hammond, K. (2014). The knowledge that 'flavours' a claim: Towards building and assessing historical knowledge on three scales. *Teaching History*, 157, pp. 18–24.

Harland, J. and Kinder, K. (1997). Teachers' continuing professional development: Framing a model of outcomes. *Journal of In-Service Education,* 23(1), pp. 71–84.

Harris, A. (2001). Department improvement and school improvement: A missing link? *British Educational Research Journal*, 27(4), pp. 477–486.

Harris, A. (2004). Distributed leadership and school improvement: Leading or misleading? *Educational Management Administration & Leadership,* 32, pp. 11–26.

Hattie, J. and Timperley, H. (2007). The power of feedback. *Review of Educational Research,* 77(1), pp. 81–112.

Heath, C. and Heath, D. (2010). *Switch: How to change things when change is hard*. London: Random House.

Howe, C. and Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education,* 43(3), pp. 325–356.

Hunter, M. C. (1982). *Mastery teaching*. El Segundo, CA: Tip Publications.

Independent Teacher Workload Review Group (2016). *Eliminating unnecessary workload around marking: Report of the Independent Teacher Workload Review Group*. Department for Education.

Jacob, R., Hill, H. and Corey, D. (2017). The impact of a professional development program on teachers' mathematical knowledge for teaching, instruction, and student achievement. *Journal of Research on Educational Effectiveness,* 10(2), pp. 379–407. DOI: 10.1080/19345747.2016.1273411.

Kahneman, D. (2011). *Thinking, fast and slow*. London: Penguin.

Kahneman, D. and Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist,* 64(6), pp. 515–526.

Kalyuga, S. and Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology*, 96(3), pp. 558–568.

Kennedy, M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), pp. 945–980.

Kingston, N. and Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice,* 30(4), pp. 28–37.

Kini, T. and Podolsky, A. (2016). *Does teaching experience increase teacher effectiveness? A review of the research*. Palo Alto: Learning Policy Institute.

Kirby, J. (2014). What can we learn from Dylan Wiliam and AfL? *Pragmatic reform* [blog], 30th March. https://pragmaticreform.wordpress.com/2013/03/30/afl

Kirschner, P. and van Merriënboer, J. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist,* 48(3), pp. 169–183.

Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Kluger, A. and DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), pp. 254–284.

Koriat, A. (2007). Metacognition and consciousness. In Zelazo, P. , Moscovitch, M. and Thompson, E. (eds.), *The Cambridge handbook of consciousness*. Cambridge: Cambridge University Press, Ch. 11.

Kraft, M. and Blazar, D. (2016). Individualized coaching to improve teacher practice across grades and subjects: New experimental evidence. *Educational Policy,* 31(7), pp. 1033–1068.

Kraft, M. and Papay, J. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*, 36(4), pp. 476–500.

Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology,* 77(6), pp. 1121–1134.

Kyun, S., Kalyuga, S. and Sweller, J. (2013). The effect of worked examples when learning to write essays in English literature. *The Journal of Experimental Education,* 81(3), pp. 385–408.

Larkin, J., McDermott, J., Simon, D. and Simon, H. (1980). Expert and novice performance in solving physics problems. *Science*, 208(4450), pp. 1335–1342.

Lemov, D. (2015). *Teach like a champion 2.0*. San Francisco, CA: Jossey-Bass.

Lemov, D., Driggs, C. and Woolway, E. (2016). *Reading reconsidered: A practical guide to rigorous literacy instruction*. San Francisco, CA: Jossey-Bass.

Lin-Siegler, X., Shaenfield, D. and Elder, A. (2015). Contrasting case instruction can improve selfassessment of writing. *Educational Technology Research and Development,* 63, pp. 1–21.

Lipsky, M. (1980[2010]). *Street-level bureaucracy: Dilemmas of the individual in public services*. New York, NY: Russell Sage Foundation.

Livingston, C. and Borko, H. (1989). Expert-novice differences in teaching: A cognitive analysis and implications for teacher education. *Journal of Teacher Education,* 37, pp. 36–42.

Locke, E. A. and Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. *American Psychologist,* 57, pp. 705–717.

Loughran, J., Berry, A. and Mulhall, P. (2012). *Understanding and developing science teachers' pedagogical content knowledge: 2nd edition*. Rotterdam: Sense Publishers.

Lu, J., Quoidbach, J., Gino, F., Chakroff, A., Maddux, W. and Galinsky, A. (2017). The dark side of going abroad: How broad foreign experiences increase immoral behaviour. *Journal of Personality and Social Psychology*, 112(1), pp. 1–16.

Marshall, B. and Drummond, M. (2006). How teachers engage with Assessment for Learning: Lessons from the classroom. *Research Papers in Education,* 21(2), pp. 133–149.

Massey, C. (2016). Asking year 12, 'what would Figes do?' Using an academic historian as the gold standard for feedback. *Teaching History,* 164, pp. 29–37.

McInerney, L. (2013). The quick, the weird, and the thorough: How I mark student work. *LauraMcInerney.com* [blog], 4th November. https://lauramcinerney.com/2013/11/04/the-quick-the-weird--and-the-thoroughhow-i-mark-student-work/

McLaughlin, M. and Talbert, J. (2001). *Professional communities and the work of high school teaching*. Chicago: University of Chicago Press.

Meyer, J. and Land, R. (2003). Threshold concepts and troublesome knowledge 1: Linkages to ways of thinking and practising. In Rust, C. (ed.), *Improving student learning: ten years on*. Oxford: OCSLD.

Millar, R. (2016). Using assessment to drive the development of teaching-learning sequences. In Lavonen, J., Juuti, K., Lampiselkä, J., Uitto, A. and Hahl, K. (eds.), *Electronic proceedings of the ESERA 2015 conference: Science education research: Engaging learners for a sustainable future,*

Part 11 (co-ed. Dolin, J. and Kind, P.). Helsinki, Finland: University of Helsinki, pp. 1631–1642. www.esera.org/publications/esera-conference-proceedings/science-education-research-esera-2015/

Millar, R. and Hames, V. (2003). *Using diagnostic assessment to enhance teaching and learning: A study of the impact of research-informed teaching materials on science teachers' practices*. Evidence-Based Practice in Science Education (EPSE) Research Network.

Mittler, P. (1973). Purposes and principles of assessment. In Mittler, P. (ed.), *Assessment for Learning in the mentally handicapped*. Edinburgh, UK: Churchill Livingstone, pp. 1–16.

Morgan, J. (n.d.). Misconceptions. *Resourceaholic* [blog]. www.resourceaholic.com/p/misconceptions.html

Muller, D. (2011). The key to effective educational science videos. *TED*. https://tedxsydney.com/talk/derek-muller-the-key-to-effective-educational-science-videos/National Trust (n.d.). *50 things to do before you're 11 ¾*. www.nationaltrust.org.uk/documents/50-thingsactivity-list.pdf

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist,* 22(2), pp. 155–175.

NCEE (2016). *Does content-focused teacher professional development work? Findings from three institute of education sciences studies*. NCEE Evaluation Brief. National Centre for Education Evaluation and Regional Assistance.

Nelson, M. and Schunn, C. (2008). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science,* 37(4), pp. 375–401.

Nuthall, G. (2007). *The hidden lives of learners*. Wellington, NZ: New Zealand Council for Educational Research.

Parkes, J. and Zimmaro, D. (2016). *Learning and assessing with multiple-choice questions in college classrooms*. New York: Routledge.

Pashler, H., Bain, P. , Bottge, B., Graesser, A., Koedinger, K., McDaniel, M. and Metcalfe, J. (2007). *Organizing instruction and study to improve student learning* (NCER 2004–2007). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Perkins, D. and Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher,* 18(1), pp. 16–25.

Pershan, M. (2017). Teaching, in general. *Teaching problems* [blog], 15th November. https://problemproblems.wordpress.com/2017/11/15/teaching-in-general/

Petrosino, A., Turpin-Petrosino, C., Hollis-Peel, M. E. and Lavenberg, J. G. (2013). 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane Database of Systematic Reviews,* DOI: 10.1002/14651858.CD002796.pub2.

Phillips, R. (2001). Making history curious: Using Initial Stimulus Material (ISM) to promote enquiry, thinking and literacy. *Teaching History,* 105, pp. 19–25.

Polanyi, M. (1962). *Personal knowledge: Towards a post-critical philosophy*. Abingdon: Routledge.

Popp, J. and Goldman, S. (2016). Knowledge building in teacher professional learning communities: Focus of meeting matters. *Teaching and Teacher Education,* 59, pp. 347–359.

Posner, G., Strike, K., Hewson, P. and Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education,* 66(2), pp. 211–227.

Potvin, P., Sauriol, É. and Riopel, M. (2015). Experimental evidence of the superiority of the prevalence model of conceptual change over the classical models and repetition. *Journal of Research in Science Teaching,* 52(8), pp. 1082–1108.

Pryor, J. and Crossouard, B. (2010). Challenging formative assessment: Disciplinary spaces and identities. *Assessment & Evaluation in Higher Education,* 35(3), pp. 265–276.

Quigley, A. (2017). The problem with past exam papers. *The confident teacher* [blog], 8th April. www.theconfidentteacher.com/2017/04/the-problem-with-past-exam-papers/

Recht, D. R. and Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology,* 80, pp. 16–20.

Renkl, A., Hilbert, T. and Schworm, S. (2009). Example-based learning in heuristic domains: A cognitive load theory account. *Educational Psychology Review,* 21(67), DOI: 10.1007/s10648-008-9093-4.

Ritter, S., Damian, R., Simonton, D., van Baaren, R., Strick, M., Derks, J. and Dijksterhuis, A. (2012). Diversifying experiences enhance cognitive flexibility. *Journal of Experimental Social Psychology,* 48(4), pp. 961–964.

Robinson, V., Lloyd, C. and Roqe, K. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly,* 44(5), pp. 635–674.

Rodriguez, M. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice,* 24(2), pp. 3–13.

Rohrer, D. and Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, 35, pp. 481–498.

Ross, J. and Bruce, C. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *The Journal of Educational Research*, 101(1), pp. 50–60.

Royal Society of Chemistry (n.d.). *Learn chemistry* [website], www.rsc.org/Learn-Chemistry

Rust, C., Price, M. and O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education,* 28, pp. 147–164.

Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), pp. 119–144.

Schmidt, H. and Rikers, R. (2007). How expertise develops in medicine: Knowledge encapsulation and illness script formation. *Medical Education,* 41, pp. 1133–1139.

Scriven, M. (1967). The methodology of evaluation. In Tyler, R. W., Gagne, R. M. and Scriven, M. (eds.), *Perspectives of curriculum evaluation*. Chicago, IL: Rand McNally, pp. 39–83.

Sealy, C. (2017). The 3D curriculum that promotes remembering. *Primerytimerydotcom* [blog], 28th October. https://primarytimery.com/2017/10/28/the-3d-curriculum-that-promotes-remembering/

Sergiovanni, T. (2005). *Strengthening the heartbeat: Leading and learning together in schools*. San Francisco: Jossey-Bass.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher,* 27(2), pp. 4–13.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), pp. 1–23.

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), pp. 153–189.

Simon, H. and Chase, W. (1973). Skill in chess. *American Scientist,* 61(4), pp. 394–403.

Smith, E. and Gorard, S. (2005). 'They don't give us our marks': The role of formative feedback in student progress. *Assessment in Education,* 12(1), pp. 21–38.

Smith, M. (2017). Cognitive validity: Can multiple-choice items tap historical thinking processes? *American Educational Research Journal*, 54(5), pp. 1256–1287. DOI: 10.3102/0002831217717949.

Soderstrom, N. and Bjork, R. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science,* 10(2), pp. 176–199.

Spillane, J. (2005). Primary school leadership practice: How the subject matters. *School Leadership & Management*, 25(94), pp. 383–397.

Stein, M., Engle, R., Smith, M. and Hughes, E. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning,* 10(4), pp. 313–340.

Stobart, G. and Stoll, L. (2005). The key stage 3 strategy: What kind of reform is this? *Cambridge Journal of Education,* 35(2), pp. 225–238. DOI: 10.1080/03057640500146906

Strachan, S. (2017). Why I love … whole class feedback & other time-saving feedback strategies. *Susanenglish* [blog], 23rd October. https://susansenglish.wordpress.com/2017/10/23/why-i-lovewhole-class-feedback-other-time-saving-feedback-strategies/

Strong, M., Gargani, J. and Hacifazlioğlu, O. (2011). Do we know a successful teacher when we see one? Experiments in the identification of effective teachers. *Journal of Teacher Education,* 62(4), pp. 367–382.

Supovitz, J. (2013). *The linking study: An experiment to strengthen teachers' engagement with data on teaching and learning*. CPRE Working Papers.

Swaffield, S. (2009). *The misrepresentation of Assessment for Learning – and the woeful waste of a wonderful opportunity*. AAIA National Conference, Bournemouth.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science,* 12, pp. 257–285.

Sweller, J., Ayres, P., Kalyuga, S. and Chandler, P. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), pp. 23–31.

Sweller, J., van Merriënboer, J. and Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review,* 10, pp. 251–296.

Tharby, A. (2017). Connecting and organising knowledge in English literature. *Reflecting English* [blog], 23rd September. https://reflectingenglish.wordpress.com/2017/09/23/connecting-and--organisingknowledge-in-english-literature/

Timperley, H. (2008). Teacher professional learning and development. *Educational Practices,* 18. International Academy of Education.

Treadaway, M. (2015). Why measuring pupil progress involves more than taking a straight line. *Education datalab* [blog], 5th March. https://educationdatalab.org.uk/2015/03/why-measuring--pupil-progressinvolves-more-than-taking-a-straight-line/

van de Grift, W. (2013). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement,* 25(3), pp. 295–311.

van de Grift, W., Helms-Lorenz, M. and Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation,* 43, pp. 150–159.

van Merriënboer, J. and Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review,* 17(2), pp. 147–177.

Vescio, V., Ross, D. and Adams, A. (2007). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education,* 24, pp. 80–91.

Wayne, A., Yoon, K., Zhu, P., Cronen, S. and Garet, M. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher,* 37(8), pp. 469–479.

Webb, T. and Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin,* 132(2), pp. 249–268. DOI: 10.1037/0033-2909.132.2.249.

Wenger, E. (2000). Communities of practice and social learning systems. *Organization,* 7, pp. 225–247.

Westerman, D. (1991). Expert and novice teacher decision making. *Journal of Teacher Education,* 42(4), pp. 292–305.

White, B. and Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction,* 16(1), pp. 3–118.

Whitehead, A. (1911). *An introduction to mathematics*. New York: Henry Holt.

Wiliam, D. (2007). Content then process: Teacher learning communities in the service of formative assessment. In Reeves, D. B. (ed.), *Ahead of the curve: The power of assessment to transform teaching and learning*. Bloomington, IN: Solution Tree, pp. 183–204.

Wiliam, D. (2010). What counts as evidence of educational achievement? The role of constructs in the pursuit of equity in assessment. *Review of Research in Education,* 34, pp. 254–284.

Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree.

Wiliam, D. (2013a). Assessment: The bridge between teaching and learning. Voices from the Middle. National Council of Teachers of English. 21(2). pp. 15–20.

Wiliam, D. (2013b). Formative assessment in mathematics: Opportunities and challenges, presented to Teachers College, Columbia University, New York, October 2013. www.dylanwiliam.org/Dylan_Wiliams_website/Presentations.html

Wiliam, D. (2016). *Leadership for teacher learning: Creating a culture where all teachers improve so*

that all students succeed. West Palm Beach, FL: Learning Sciences International.

Wiliam, D. (2017). Assessment, marking and feedback. In Hendrick, C. and McPherson, R. (eds.), *What does this look like in the classroom? Bridging the gap between research and practice*. Woodbridge: John Catt.

Wiliam, D. (2018). *Creating the schools our children need: Why what we're doing right now won't help much and what we can do instead*. Learning Sciences International.

Wiliam, D. and Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal,* 22(5), pp. 537–548.

Wiliam, D. and Leahy, S. (2014). *Sustaining formative assessment with teacher learning communities*. Learning Sciences, Dylan Wiliam Center.

Willingham, D. (2004). Ask the cognitive scientist: The privileged status of story. *American Educator,* Summer.

Willingham, D. (2009). *Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom*. San Francisco, CA: Jossey-Bass.

Wittwer, J. and Renkl, A. (2010). How effective are instructional explanations in example-based learning? A meta-analytic review. *Educational Psychology Review,* 22(4), pp. 393–409.

Wood, W. and Neal, D. (2007). A new look at habits and the habit: Goal interface. *Psychological Review,* 114(4), pp. 843–863.

Wylie, C. and Wiliam, D. (2007). Analyzing diagnostic items: What makes a student response interpretable?, paper presented at the annual meeting of the National Council on Measurement in Education (NCME), Chicago, IL, April 2006. www.dylanwiliam.org/Dylan_Wiliams_website/Papers_fi les/DIMS%20%28NCME%202007%29.pdf

Yeager, D., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P. , Master, A., Hessert, W., Williams, M. and Cohen, G. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General,* 143(2), pp. 804–824.

Yeager, D. and Walton, G. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research,* 81(2), pp. 267–301.

Yin, Y., Shavelson, R., Ayala, C., Ruiz-Primo, M., Brandon, P., Furtak, E., Tomita, M. and Young, D. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education,* 21(4), pp. 335–359.

Young, M. (2014a). The progressive case for a subject-based curriculum. In Young, M., Lambert, D., Roberts, C. and Roberts, M. (eds.), *Knowledge and the future school: Curriculum and social justice*. London: Bloomsbury, Ch. 4.

Young, M. (2014b). Knowledge, curriculum and the future school. In Young, M., Lambert, D., Roberts, C. and Roberts, M. (eds.), *Knowledge and the future school: Curriculum and social justice*. London: Bloomsbury, Ch. 1.

Zhu, X. and Simon, H. (1987). Learning mathematics from examples and by doing. *Cognitive Instruction,* 4, pp. 137–166.

Zimmerman, B. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice,* 41(2), pp. 64–70.