Principles for Assessment Design and Use in the Service of Learning

This page outlines principles that guide the design and use of learning-focused assessments intended to support student learning. In the Handbook volumes, the principles were intended to assist chapter authors in considering these common elements in their contributions.

- **Principle 1:** Assessment **transparency** provides clear information about assessment content and use to assist learners, teachers, administrators, and parents.
- **Principle 2:** Assessment **focus** is explicit and includes purposes, outcomes, progress indicators, and processes that can be transferred to other settings, situations, and conditions.
- **Principle 3:** Assessment design supports learners' **processes**, such as motivation, attention, engagement, effort, and metacognition.
- **Principle 4:** Assessments model the structure of **expectations** and **desired learning** over time.
- **Principle 5: Feedback**, adaptation, and other relevant instruction should be linked to assessment experiences.
- **Principle 6:** Assessment **equity** requires fairness in design of tasks and their adaptation to permit their use with respondents of different backgrounds, knowledge, and experiences.
- **Principle 7:** Assessment **quality and validity** should be available and reflect evidence related to assessment purpose to permit appropriate inferences and findings about quality, utility, and credibility.

For those interested in the scientific or experiential bases of the principles, we refer you to the selected bibliography below. For each principle, the selected bibliography provides a set of references that highlight its theoretical and empirical underpinnings.

For more information, please refer to:

Baker, E. L., Everson, H. T., Tucker, E. M., & Gordon, E. W. (2025). Principles for assessment in the service of learning. In E. M. Tucker, E. Armour-Thomas, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning, Volume I: Foundations for Assessment in the Service of Learning.* University of Massachusetts Amherst Libraries.

Selected Bibliography

Assessment in the Service of Learning

Baker, E. L., & Gordon, E. W. (2014). From the assessment of education to the assessment for education: Policy and futures. *Teachers College Record, 116,* 1–24.

Darling-Hammond, L., & Adamson, F. (2014). *Beyond the bubble test: How performance assessments support 21st-century learning.* Jossey-Bass.

Gordon, E. W. (2020). Toward assessment in the service of learning. *Educational Measurement: Issues and Practice, 39*(3), 72–78.

The Gordon Commission on the Future of Assessment in Education. (2013). *To assess, to teach, to learn: A vision for the future of assessment* [Technical Report]. https://www.ets.org/Media/Research/pdf/gordon_commission_technical_report.pdf

Pellegrino, J. (2014). Assessment in the service of teaching and learning: Changes in practice enabled by recommended changes in policy. *Teachers College Record, 116*(110313). https://doi.org/10.1177/016146811411601102

Ruiz-Primo, M. A., & Furtak, E. M. (2024). Classroom activity systems to supportambitious teaching and assessment. In S. F. Marion, J. W. Pellegrino, & A. I. Berman (Eds.), *Reimagining balanced assessment systems* (pp. 93–131). National Academy of Education.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Principle 1: Assessment transparency provides clear information about assessment content and use to assist learners, teachers, administrators, and parents.

Chung, G. K. W. K., Delacruz, G. C., Dionne, G. B., & Bewley, W. L. (2003). Linking assessment and instruction using ontologies. *Proceedings of the I/ITSEC, 25,* 1811–1822.

Clancey, W. J., & Shortliffe, E. H. (Eds.). (1984). *Readings in medical artificial intelligence: The first decade.* Addison-Wesley. https://impact.dbmi.columbia.edu/~ehs7001/Clancey-Shortliffe-1984/Readings%20Book.htm

Gagné, R. M., & Briggs, L. J. (1974). *Principles of instructional design.* Holt, Rinehart & Winston.

Iseli, M. R., & Jha, R. (2016). Computational issues in modeling user behavior in serious games. In H. F. O'Neil, E. L. Baker, & R. S. Perez (Eds.), *Using games and simulations for teaching and assessment: Key issues* (pp. 21–40). Routledge.

Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education, 39*(7), 840–852. https://doi.org/10.1080/02602938.2013.875117

Moss, C. M., & Brookhart, S. M. (2012). *Learning targets: Helping students aim for understanding in today's lesson.* ASCD.

Principle 2: Assessment focus is explicit and includes purposes, outcomes, progress indicators, and processes that can be transferred to other settings, situations, and conditions.

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives* (Complete edition). Longman.

Andrade, H. L., Bennett, R. E., & Cizek, G. J. (Eds.). (2019). *Handbook of formative assessment in the disciplines* (1st ed.). Routledge. https://doi.org/10.4324/9781315166933

Armour-Thomas, E., & Gordon, E. W. (2025). *Principles of dynamic pedagogy: An integrative model of curriculum instruction and assessment for prospective and in-service teachers.* Routledge.

Chatterji, M. (2025). *User-centered assessment design: An integrated methodology for diverse populations.* Guilford Press.

Heritage, M. (2021). *Formative assessment: Making it happen in the classroom* (2nd ed.). Corwin.

Lee, C. D. (1998). Culturally responsive pedagogy and performance-based assessment. *The Journal of Negro Education, 67*(3), 268–279. https://www.jstor.org/stable/2668195?origin=crossref

van Merriënboer, J. J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning: A systematic approach to four-component instructional design.* Routledge.

Principle 3: Assessment design supports learners' processes, such as motivation, attention, engagement, effort, and metacognition.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school* (Expanded ed.). National Academy Press.

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

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066x.34.10.906

Plass, J. L., & Kalyuga, S. (2019). Four ways of considering emotion in Cognitive Load Theory. *Educational Psychology Review, 31*(2), 339–359. https://doi.org/10.1007/s10648-019-09473-5

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Principle 4: Assessments Model the structure of expectations and desired learning over time.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Black, P., Wilson, M., & Yao, S.-Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research and Perspectives, 9*(2–3), 71–123.

Darling-Hammond, L., Herman, J., Pellegrino, J. W., Abedi, J., Aber, J. L., Baker, E., Bennett, R., Gordon, E., Haertel, E., Hakuta, K., Ho, A., Linn, R. L., Pearson, P. D., Popham, W. J., Resnick, L., Schoenfeld, A. H., Shavelson, R., Shepard, L. A., Shulman, L., & Steele, C. M. (2013). *Criteria for high-quality assessment.* Stanford Center for Opportunity Policy in Education.

Gordon, E. G., & Bridgall B. L. (Eds.). (2006). *Affirmative development: Cultivating academic ability, critical issues in contemporary American education series*. Rowman & Littlefield Publishers, Inc.

Leonard, W. H., & Lowery, L. F. (1984). The effects of question types in textual reading upon retention of biology concepts. *Journal of Research in Science Teaching, 21*(4), 377–384. https://doi.org/10.1002/tea.3660210405

Phelps, R. P. (2012). The effects of testing on student achievement, 1910–2010. *International Journal of Testing, 12,* 21–43.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning. *Psychological Science, 17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Principle 5: Feedback, adaptation, and other relevant instruction should be linked to assessment experiences.

Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement* (1st ed.). Routledge.

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

Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*(1), 79–97. https://doi.org/10.3102/00346543058001079

Linn, R. L. (1983). Testing and instruction: Links and distinctions. *Journal of Educational Measurement, 20*(2), 179–189.

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

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10,* Article 3087. https://doi.org/10.3389/fpsyg.2019.03087

Principle 6: Assessment equity requires fairness in design of tasks and their adaptation to permit their use with respondents of different backgrounds, knowledge, and experiences.

Armour-Thomas, E., McCallister, C., Boykin, A. W., & Gordon, E. W. (Eds.). (2019). *Human variance and assessment for learning.* Third World Press.

Bennett, R. E. (2023). *T*oward a theory of socioculturally responsive assessment. *Educational Assessment, 28*(2), 83–104. https://doi.org/10.1080/10627197.2023.2202312

Duran, R. P. (1989). Testing of linguistic minorities. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 573–587). Macmillan.

Gordon, E. W. (1995). Toward an equitable system of educational assessment. *The Journal of Negro Education, 64*(3), 360–372.

Herman, J. L., Bailey, A. L., & Martinez, J. F. (2023). Introduction to the special issue: Fairness in educational assessment and the next edition of the standards. *Educational Assessment, 28*(2), 65–67. https://doi.org/10.1080/10627197.2023.2215979

Nasir, N. S., Lee, C. D., Pea, R., & McKinney de Royston, M. (Eds.). (2020). *Handbook of the cultural foundations of learning.* Routledge.

Oakes, J. (1986). Keeping track, part 1: The policy and practice of curriculum inequality. *Phi Delta Kappan, 68*(1), 12–17. https://www.jstor.org/stable/20403250

Shepard, L. A. (2021). Ambitious teaching and equitable assessment: A vision for prioritizing learning, not testing. *American Educator, 45*(3), 28–37, 48.

Sireci, S. G., & Randall, J. (2021). *Evolving notions of fairness in testing in the United States in the history of educational measurement.* Routledge.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher, 32*(2), 3–13. https://doi.org/10.3102/0013189x032002003

Principle 7: Quality and validity should be available and reflect evidence related to assessment purpose to permit appropriate inferences and findings about quality, utility, and credibility.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* American Educational Research Association.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73.

Linn, R. L. (2010). Validity. In B. McGaw, P. L. Peterson, & E. L. Baker (Eds.), *International encyclopedia of education* (3rd ed., Vol. 4, pp. 181–185). Elsevier.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). American Council on Education/Macmillan.

Mislevy, R. J., Oliveri, M. E., Slomp, D., Crop Eared Wolf, A., & Elliot, N. (2025). An evidentiary-reasoning lens for socioculturally responsive assessment. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 199–241). Routledge/Taylor & Francis.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*(1), 3–67.

Pellegrino, J. W., DiBello, L. V., & Goldman, S. R. (2016). A framework for conceptualizing and evaluating the validity of instructionally relevant assessments. *Educational Psychologist, 51*(1), 59–81.

Sireci, S. G. (2013). Agreeing on validity arguments. *Journal of Educational Measurement, 50*(1), 99–104.