Standard 12: Assessment and Data

The program trains teacher candidates how to assess learning and use student performance data to inform instruction.

Why this standard?

Great teachers know what concepts and skills their students have mastered and what they still have trouble understanding. Not surprisingly, there is increasing evidence that the better a teacher becomes at pinpointing what students need to learn, the better the students do learn. The capacity to develop assessments (e.g., quizzes), as well as to analyze and interpret assessment results to improve instruction, are explicit skills teacher candidates should have an opportunity to practice.

What is the focus of the standard?

Coursework and assignments representing the culmination of a candidate's preparation are examined to check that elementary and secondary teacher candidates have an opportunity to practice developing their own assessments, analyzing student assessment results, and applying their analysis to lesson planning. We also check to see that candidates have an opportunity to practice analyzing student data in teams because schools are increasingly fostering a collaborative approach to teaching. The "strong design" (🏆) indicator evaluates the structure of coursework in which preparation related to assessment is provided, examining whether a core assessment course is complemented by coverage of subject-specific assessment topics in methods courses.

Standard applies to elementary and secondary programs.

Standard and Indicators

Standard 12: Assessment and Data

The program trains teacher candidates how to assess learning and use student performance data to inform instruction.
Standard applies to: Elementary and Secondary programs.

Indicators that the program meets the standard:

12.1 The instructional role of standardized tests, particularly the program state's standardized tests, is reviewed by teacher candidates.

12.2 Teacher candidates are required to prepare formative and summative classroom assessments.

12.3 Individually and in teams, teacher candidates are taught how to interpret and apply data from both standardized and classroom assessments in order to inform instruction.

STRONG DESIGN

12.4 A program will receive a "strong design" designation if instruction satisfying indicators 12.1-12.3 is provided through a combination of 1) a core data literacy course also addressing the use and misuse of assessment data, the issue of bias, and the meaning of validity and reliability in the context of testing, and 2) one or more courses addressing subject-specific pedagogical data literacy and including field work assessment assignments closely aligned with that instruction.

Rationale
Standard 12: Assessment and Data

The program trains teacher candidates how to assess learning and use student performance data to inform instruction.

Standard applies to elementary and secondary programs.

Why this standard?

Great teachers know what concepts and skills their students have mastered and what they still have trouble understanding. Not surprisingly, there is increasing evidence that the better a teacher becomes at pinpointing what students need to learn, the better the students *do* learn. The development of assessments (e.g., quizzes), as well as the capacity to analyze and interpret assessment results to improve instruction, are explicit skills teacher candidates should have an opportunity to practice.

What is the focus of the standard?

Coursework and assignments representing the culmination of a candidate's preparation are examined to check that elementary and secondary teacher candidates have an opportunity to practice developing their own assessments, analyzing student assessment results, and applying their analysis to lesson planning. We also check to see that candidates have an opportunity to practice analyzing student data in teams because schools are increasingly fostering a collaborative approach to teaching. The "strong design" (🏆) indicator evaluates the structure of coursework in which preparation related to assessment is provided, examining whether a core assessment course is complemented by coverage of subject-specific assessment topics in methods courses.

Rationale

Research base for this standard

No "strong research"[1] exists with regard to assessment and efficacy of using data for teacher preparation.

However, the application of assessments and data can be beneficial within K-12 education. Some additional research[2] has found that taking an assessment can be beneficial to students as it enhances learning through strengthened memory representation.[3] One study found that frequent quizzes during a lecture decreased

---

[1] NCTQ has created "research inventories" that describe research conducted within the last decade or so that has *general* relevance to aspects of teacher preparation also addressed by one or more of its standards (with the exceptions of the Outcomes and Evidence of Effectiveness standards). These inventories categorize research along two dimensions: design methodology and use of student performance data. Research that satisfies our standards on both is designated as "strong research" and will be identified as such. That research is cited here if it is *directly* relevant to the standard; strong research is distinguished from other research that is not included in the inventory or is not designated as "strong" in the inventory. Refer to the introduction to the research inventories for more discussion of our approach to categorizing research. If a research inventory has been developed to describe research that generally relates to the same aspect of teacher prep as addressed by a standard, the inventory can be found in the back of this standard book.

[2] "Additional research" is research that is not designated as "strong" because it is not as recent and/or does not meet the highest standards for design methodology and/or use of student performance data.

[3] Chang, C. Y., Yeh, T. K., & Barufaldi, J. P. (2010). The positive and negative effects of science concept tests on student conceptual understanding. *International Journal of Science Education, 32*(2), 265-282; Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science, 331*, 772-775; McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger III, H. L. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology, 103*(2), 399-414; Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Education Researcher, 39*(5), 406-412; Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 36*(1), 233-239; Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teach about K-12 assessment: A review.* Retrieved February 12, 2013, from http://www.nctq.org/p/publications/docs/assessment_report.pdf

students' mind wandering, increased note-taking, increased performance or a cummulative test, and reduced anxiety about that final test.[4]

Other support for this standard
Emerging evidence from highly effective schools suggests that making instruction more "data-driven" (i.e., using assessment data to adjust what teachers do) can play a significant role in teacher effectiveness.[5] A recent survey of school staff found that over one-quarter (29 percent) of new teachers were not at all prepared or only somewhat prepared to assess students.[6]

On the basis of emerging research findings, the Institute of Education Sciences provides five recommendations for how educators should use data. These recommendations specify that teachers should use assessment information in a cyclical manner whereby they collect data about student learning, interpret those data to determine how to improve students' learning, modify instruction based on their determination and collect new data.[7] To successfully implement this cycle, teachers must be able to implement both formative and summative strategies to assess their students' learning, to analyze assessment outcome data and to adjust instruction based on those data.

This standard also receives support from school district superintendents.

---

[4] Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences, 110*(16), 6313-6317. While this study focuses on college students in an online class, the results may still be generalizable to K-12 students in a classroom setting.

[5] Bambrick-Santoyo, P. (2010). *Driven by data: A practical guide to improve instruction.* San Francisco: Jossey-Bass.

[6] Coggshall, J. G., Bivona, L., & Reschly, D. J. (2012, August). *Evaluating the effectiveness of teacher preparation programs for support and accountability*. Washington, DC: National Comprehensive Center for Teacher Quality.

[7] Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). Using student achievement data to support instructional decision making (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved February 11, 2013, from http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Scoring Methodology
How NCTQ scores the Assessment and Data Standard

[Standard and indicators](#)

Data used to score this standard

Evaluation of elementary and secondary programs on Standard 12: Assessment and Data uses the following sources of data:

- Syllabi of relevant required courses

- Capstone projects—often called teacher work samples (TWS) or portfolios and including teacher performance assessments (TPAs)[1]

Who analyzes the data

Two [general analysts](#) independently evaluate each elementary and secondary program using a detailed scoring protocol from which this scoring methodology is abstracted. For information on the process by which scoring discrepancies are resolved, see the "scoring processes" section of the [General Methodology](#).

Scope of Analysis

For elementary preparation programs, this analysis is based on evaluation of syllabi in required courses whose descriptions mention assessment[2] and the syllabus for any required math methods course if not already included. Syllabi of required literacy courses are excluded from analysis.[3] For secondary preparation programs, analysis is based on evaluations of syllabi in required courses whose descriptions mention assessment, as well as the syllabus for the required methods course(s) in one core subject.[4] In analysis of both elementary and

---

[1] Analysts evaluate any assessment-related project assignments, whether these projects require candidates to submit artifacts from earlier coursework that reflect assessment assignments or entail new assessment assignments.

In **California**, **Minnesota** and **Washington**, the relevant TPA used in the state is utilized as a data source for all IHEs regardless of whether provided by the IHE because TPA initiatives are statewide. In **Tennessee** and **Ohio**, the relevant TPA used by the state is used as a data source for selected IHEs regardless of whether provided by the IHE if a state official indicated that the IHE has moved beyond isolated field tests to a full implementation. In all other states, a TPA is used as a data source only if provided by the IHE.

In many cases, the TPA used by the IHE is the "edTPA," a proprietary instrument that NCTQ could not obtain in its entirety for a full evaluation. However, we were able to obtain a portion of the edTPA and compare it to the Stanford PACT (the edTPA's precursor), and we were also able to confirm with a knowledgeable state education agency official that the edTPA does not substantively differ from the Stanford PACT in any of the features on which we base our evaluation.

In **Oregon**, the state's required Teacher Work Sample is utilized as a data source.

[2] Terms in titles or coursework that are considered to indicate coverage of assessment are "assessment," "measurement," "evaluation," "evaluate instruction," "monitoring of student progress," "evaluation of learning," "analysis of outcomes" and "instructional planning tied to/based on assessment."

[3] The more specialized types of assessment emphasized in literacy courses fulfill different purposes and require a more specialized evaluation than provided in this standard. An analysis of assessment in literacy courses is conducted as part of the evaluation of **Standard 2: Early Reading**.

[4] To parallel the experience of secondary candidates, we examine a randomly selected pathway for certification (mathematics, the sciences, English, or the social sciences) and then evaluate the methods course for the pathway chosen, regardless of whether assessment is mentioned in the course's title or description. The same pathway is used for our evaluation of **Standard 15: Secondary Methods**.

secondary coursework, educational psychology courses taught in a psychology department and special education courses are excluded from analysis.[5]

In both elementary and secondary programs, capstone projects are also evaluated. Thorough document processing ensures that capstone project descriptions and any other relevant sources of data from student teaching are isolated.

In both **undergraduate** and **graduate** programs, Indicator 12.1 is evaluated for reporting purpose (not scoring) using syllabus descriptions of course lecture topics and assigned reading (with course objectives and assignment descriptions providing context as necessary for interpretation). For evaluation of this indicator, references to the state's standardized tests found in a non-pedagogical context (e.g., in a "Foundations of Education" course) do not satisfy the indicator.

Indicators 12.2 and 12.3 are evaluated for scoring purposes using assignments noted in syllabi and descriptions of capstone projects. Throughout the evaluation, evidence may be taken from several lectures or assignments within one course (and a capstone project, as relevant) as well as across several courses (and the capstone project, as relevant) in order to satisfy a particular indicator. (More discussion of analysis using syllabi.)

Analysts are trained to take the broadest possible interpretation of assessment references and also to consider standard usage. For example, a reference to "accountability" in the context of a lecture on assessment is presumed to refer to the state's standardized test system because this broad interpretation is justified. However, in the absence of any other description, a requirement that a teacher candidate develop an "authentic assessment" is presumed to satisfy the requirement that candidates develop summative assessments, not both formative and summative assessments, because "authentic assessment" most commonly is posed as an alternative to more traditional forms of summative assessment.

Common misconceptions about how analysts evaluate the Assessment and Data Standard:

■ *Syllabi are the only data source considered in the standard's evaluation.* Evaluation for this standard does beyond evaluation of syllabi to include culminating projects – which often contain the only assignments involving data analysis and interpretation required of teacher candidates.

■ *Lectures addressing the preparation of classroom assessments and interpretation of assessment data can be used for evaluation of this standard.* Hearing about assessment isn't sufficient to ensure that teacher candidates will enter the classroom with sufficient preparation: To receive credit, a program's teacher candidates must engage in practice activities or assignments that result in tangible products.

■ *Assessment related coursework in special education and literacy courses can be used for evaluation of this standard.* Instruction on the more specialized types of assessment in special education and literacy fulfills different purposes and requires a more specialized evaluation than provided in this standard. (An analysis of assessment in literacy courses *is* conducted as part of the evaluation of **Standard 2: Early Reading**.)

Due to the burden imposed by document processing and analysis, the full sample of programs in the first edition of the *Review* was not evaluated on this standard. Instead of evaluating all programs for which any material had been provided by an IHE, we instead established a calendar deadline for analysis that would ensure that we could evaluate a sample of sufficient size to provided credible information about the nature of teacher preparation in

---

[5] Treatment of assessment in educational psychology courses taught in a psychology department that are not designed for teachers are not sufficiently contextualized to K-12 schooling. Special education coursework requires a more specialized evaluation of assessment topics than provided here.

this area. Once this deadline was established, we also prioritized evaluation of programs producing the largest number of teacher candidates each year. Programs were only removed from the sample if it was impossible to make a determination on one or more indicators due to a lack of clarity in data.

How a program earns a "strong design" rating

Evaluation for strong design under this standard entails examination of the structure of assessment-relevant coursework to determine if a core data literacy course is complemented by courses that address assessment through the lens of subject-specific pedagogy and include aligned field work.

Examples of what satisfies or does not satisfy the standard's indicators

Instructional role of state's standardized tests (Indicator 12.1)

✔ · fully satisfies the indicator (reported only)	✘ · does not satisfy the indicator (reported only)
The program explicitly addresses the instructional role of standardized tests, particularly the program state's standardized tests.	The program does not explicitly address the instructional role of standardized tests, or it addresses standardized tests other than the state's tests.

Examples:

- *Lecture entitled "**Tennessee** Value-Added Assessment System."*

- *Lecture on standardized testing with references to the role of assessment in light of No Child Left Behind legislation.*

- *Lecture entitled "Standardized Tests and Accountability."*

- *Capstone project that requires teacher candidates to report and reflect on their school's state testing data and progress toward Annual Yearly Progress (AYP).[6]*

- *Course objective outline that lists coverage of the **North Carolina** EOGs, and list of lecture topics that includes "Standardized Testing."*

- *Assignment to analyze the **Georgia** Performance Standards for grades 6-8 in candidate's primary and secondary certification areas, compare to expectations for student achievement on the CRCT and analyze how NCLB mandates affect instructional priorities.*

Examples:

- *Lecture entitled "National and International Assessment Data."*

- *Lecture entitled "Standardized Assessment" in an Educational Psychology course.*

- *Lecture entitled "Large scale tests, issues, and interpretation."*

---

[6] The specificity of this assignment (in its reference to "AYP") allows the analyst to presume that instruction in coursework prepares the candidate on the topic of the state's standardized tests even though no lecture on the topic was evident.

Preparation of formative and summative assessments (Indicator 12.2)

✔ · fully satisfies the indicator	✘ · does not satisfy the indicator
The program requires teacher candidates to prepare both formative and summative assessments. Examples: ■ *Teacher candidates prepare a unit plan using the Understanding by Design (UbD) framework.[7]* ■ *Capstone project requires a full assessment plan, including a pre/post-test and several formative assessments.* ■ *Assignment entitled Thematic Unit requires several unspecified assessments. (These are interpreted to include practice relevant to this indicator because the syllabus lists requirements that include the development of formative, summative and authentic assessments.)*	The program requires the teacher candidate to prepare either formative or summative assessments (not both), or to prepare neither. Examples: ■ *Teacher candidates are required to conduct an informal, formative interview to determine a student's baseline knowledge of fractions, but the course assignments do not include a follow-up inventory of knowledge gained post-instruction.* ■ *Capstone project requires the use of formative and summative assessments; however, the assessments need not be prepared by the teacher candidate.*

---

[7] Use of the UbD planning format presumes use of formative assessment; preparation of a unit plan is always presumed to entail preparation of a summative assessment.

Individual and team analysis and interpretation of data (Indicator 12.3)

Note that this indicator requires that teacher candidates work with both classroom and standardized assessment data, both individually *and* with their peers.

✔ · fully satisfies the indicator	~ · partly satisfies the indicator	✘ · does not satisfy the indicator
The program requires teacher candidates, individually and in teams, to analyze and interpret classroom and standardized assessment data in order to inform instruction. Example: *A course requires teacher candidates to bring in **standardized <u>and</u> classroom assessment** samples and data from their field classrooms **for in-class analysis and discussion of instructional implications.** The capstone project in the same program requires teacher candidates to conduct **individual analyses of classroom and standardized assessment data as well as reflect on the instructional implications** of those assessment results.*	The program partly satisfies this indicator because teacher candidates do some combination that constitutes the majority of the activities required by this indicator. Example: *A TWS requires that teacher candidates **individually use classroom assessment data** to analyze the impact of instruction on student learning "in terms of number of students who achieved and made progress toward each learning goal." It must also be evident that the teacher candidate used pre-assessment data to modify instruction and reflected on changes he or she might make in the future based on post-assessment data. Teacher candidates must **present to their peers their classroom assessment data analysis and implications for future instruction** for feedback and discussion.*	The program does not satisfy this indicator because teacher candidates do none or only a few of the activities required by this indicator. Examples: *A program's capstone project requires teacher candidates to individually analyze pre/post-classroom assessment data, modify instruction based on formative assessment results and discuss future instructional implications based on summative assessment data.* OR *A course's field experience requirement includes a journal in which teacher candidates reflect on assessment strategies used and changes in student achievement in their field classroom.*[8]

---

[8] Assignments counted as practice for Indicators 12.2 and 12.3 cannot simply involve a teacher candidate writing reflectively about assessments developed or assessment data analyzed; they must involve the tangible product of an assessment or the analysis itself, products that can be submitted for instructor review. However, an assignment involving a written reflection that specifically discusses the instructional implications of assessment data can be counted as a practice assignment even if there is no implementation of instruction.

Research Inventory

Researching Teacher Preparation:
Studies investigating the skills and techniques teachers need for assessing students and analyzing data

These studies address issues most relevant to Standard 12: Assessment and Data

Total Number of Studies	Studies with Stronger Design		Studies with Weaker Design	
	Measures Student Outcomes	Does Not Measure Student Outcomes	Measures Student Outcomes	Does Not Measure Student Outcomes
27	0	2 Citations: 7, 24	8 Citations: 9, 13, 14, 16, 17, 20, 21, 26	17 Citations: 1–6, 8, 10–12, 15, 18, 19, 22, 23, 25, 27

Citations for articles categorized in the table are listed below.

Databases: Education Research Complete and Education Resource Information Center (peer-reviewed listings of reports on research including United States populations).

Publication dates: Jan 2000 – June 2012

See *Research Inventories: Rationale and Methods* for more information on the development of this inventory of research.

1. Bangert, A., & Kelting-Gibson, L. (2006). Teaching principles of assessment literacy through teacher work sample methodology. *Teacher Education and Practice, 19*(3), 351–364.

2. Bellara, A. P., & Hibbard, S. T. (2010). Assessing learner needs through formative evaluations in a prescriptive course: Self-reflection of teaching practices through student input. *International Journal of Learning, 17*(7), 359–368.

3. Bennett, K., & Cunningham, A. C. (2009). Teaching formative assessment strategies to pre-service teachers: Exploring the use of handheld computing to facilitate the action research process. *Journal of Computing in Teacher Education, 25*(3), 99–105.

4. Buck, G. A., Trauth-Nare, A., & Kaftan, J. (2010). Making formative assessment discernable to pre-service teachers of science. *Journal of Research in Science Teaching, 47*(4), 402–421.

5. Campbell, C., & Collins, V. L. (2007). Identifying essential topics in general and special education introductory assessment textbooks. *Educational Measurement: Issues & Practice, 26*(1), 9–18.

6. Dass, M. (2005). Using a science/technology/society approach to prepare reform-oriented science teachers: The case of a secondary science methods course. *Issues in Teacher Education, 14*(1), 95–108.

7. DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*(4), 419–438.

8. Dorfman, A. B., Galluzzo, G. R., & Meisels, S. J. (2006). Learning to teach: Developing assessment skills when program and placement are aligned. *Journal of Early Childhood Teacher Education, 27*(3), 231–247.

9. Fan, Y., Wang, T., & Wang, K. (2011). A web-based model for developing assessment literacy of secondary in-service teachers. *Computers & Education, 57*(2), 1727–1740.

10. Forbush, D. E., Stenhoff, D. M., Vasquez III, E., Furzland, M., Alexander, M., & Stein, J. (2007). Evaluation of an online tool for assessing competence in achievement testing. *Teacher Education and Special Education, 30*(3), 142–154.

11. Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through pre-service teacher education. *Teaching and Teacher Education: An International Journal of Research and Studies, 21*(6), 607–621.

12. Huai, N., Braden, J. P., White, J. L., & Elliott, S. N. (2006). Effect of an internet-based professional development program on teachers' assessment literacy for all students. *Teacher Education and Special Education, 29*(4), 36–52.

13. Hunsaker, S. L., Nielsen, A., & Bartlett, B. (2010). Correlates of teacher practices influencing student outcomes in reading instruction for advanced readers. *Gifted Child Quarterly, 54*(4), 273–282.

14. Jie-Qi, C., & McNamee, G. (2006). Strengthening early childhood teacher preparation: Integrating assessment, curriculum development, and instructional practice in student teaching. *Journal of Early Childhood Teacher Education, 27*(2), 109–128.

15. Mayor, S. (2005). Pre-service teachers' developing perspectives on assessment and remediation of struggling readers. *Reading Improvement, 42*(3), 164.

16. Menzies, H. M., Mahdavi, J. N., & Lewis, J. L. (2008). Early intervention in reading: From research to practice. *Remedial and Special Education, 29*(2), 67–77.

17. Moore, R. (2006). Taking action: Assessing the impact of pre-service teaching on learning. *Action in Teacher Education, 28*(3), 53–60.

18. Morrison, J. A. (2005). Using science notebooks to promote pre-service teachers' understanding of formative assessment. *Issues in Teacher Education, 14*(1), 5–21.

19. Morrison, J. A., McDuffie, A., & Akerson, V. L. (2005). Pre-service teachers' development and implementation of science performance assessment tasks. *International Journal of Science and Mathematics Education, 3*(3), 379–406.

20. Ruiz-Primo, M., & Furtak, E. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment, 11*(3–4), 237–263.

21. Sato, M., Wei, R., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of national board certification. *American Educational Research Journal, 45*(3), 669–700.

22. Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Pre-service teachers' assessment literacy. *Journal of Science Teacher Education, 22*(4), 371–391.

23. Southerland, S. A., Sowell, S., & Enderle, P. (2011). Science teachers' pedagogical discontentment: Its sources and potential for change. *Journal of Science Teacher Education, 22*(5), 437–457.

24. Stobaugh, R., Tassell, J., & Norman, A. D. (2010). Improving pre-service teacher preparation through the teacher work sample: Exploring assessment and analysis of student learning. *Action in Teacher Education, 32*(1), 39–53.

25. Tomanek, D., Talanquer, V., & Novodvorsky, I. (2008). What do science teachers consider when selecting formative assessment tasks? *Journal of Research in Science Teaching, 45*(10), 1113–1130.

26. Volante, L., & Beckett, D. (2011). Formative assessment and the contemporary classroom: Synergies and tensions between research and practice. *Canadian Journal of Education, 34*(2), 239–255.

27. Windschitl, M., Thompson, J., & Braaten, M. (2011). Ambitious pedagogy by novice teachers: Who benefits from tool-supported collaborative inquiry into practice and why? *Teachers College Record, 113*(7), 1311–1360.