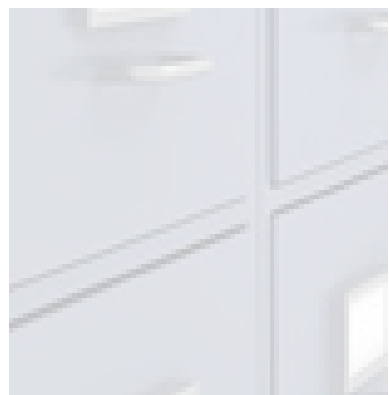


WHAT TEACHER PREPARATION PROGRAMS TEACH ABOUT K-12 ASSESSMENT: A review

May 2012



AUTHORS:

Julie Greenberg and Kate Walsh

OUR THANKS TO:

Research Analysts: Christine White, with assistance from Tom Byrne and Amy MacKown

Database Design and Technical Support: Jeff Hale

Graphic Design: Colleen Hale

PRINCIPAL FUNDING:

Michael and Susan Dell Foundation

NCTQ BOARD OF DIRECTORS:

Barbara O'Brien, Chair, Kate Walsh, President, Stacey Boyd, Chester E. Finn, Ira Fishman, Marti Watson Garlett, Henry L. Johnson, Donald N. Langenberg, Clara M. Lovett, E. Mike Miles, Paul G. Pastorek, Carol G. Peck, John L. Winn, Vice Chair

NCTQ ADVISORY BOARD:

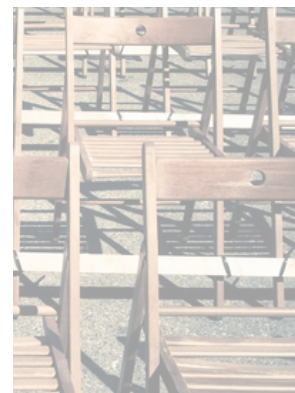
Steven J. Adamowski, Michael Barber, Roy E. Barnes, McKinley Broome, Cynthia G. Brown, David Chard, Andrew Chen, Jean Clements, Celine Coggins, Pattie Davis, Jo Lynne DeMary, Michael Feinberg, Michael Goldstein, Eric A. Hanushek, Joseph Hawkins, Frederick M. Hess, Paul T. Hill, E.D. Hirsch, Michael Johnston, Barry Kaufman, Frank Keating, Joel Klein, Martin J. Koldyke, Wendy Kopp, James Larson, Tom Lasley, Amy Jo Leonard, Deborah McGriff, Ellen Moir, Robert H. Pasternack, Michael Podgursky, Michelle Rhee, Stefanie Sanford, Audrey Soglin, Daniel Willingham, Suzanne Wilson

Preface

This report provides information on the preparation provided to teacher candidates from teacher training programs so that they can fully use assessment data to improve classroom instruction. It follows a similar shorter report we released in March 2012.¹ In this newest report, we have significantly expanded our initial sampling to encompass 180 undergraduate and graduate programs housed in 98 institutions in 30 states.²

In early 2013, with U.S. News & World Report as the publisher, we will release a comprehensive review of all teacher preparation programs in the United States. That review will also include further analysis of teachers' preparation in use of assessment for instructional improvement.

Our purpose with both preliminary reports is to encourage teacher preparation programs, along with the federal and state agencies supporting and overseeing these programs, to focus more on the importance of future teachers' knowledge and skills in the increasingly critical area of assessment.



1 http://www.nctq.org/p/edschools/docs/assessment_publication.pdf

2 Institutions included in the sample are listed in Appendix 1. Of the 180 programs, 87 (48 percent) are elementary (84 percent undergraduate and 16 percent graduate) and 93 (52 percent) are secondary (74 percent undergraduate and 26 percent graduate).





Introduction

Why teacher knowledge of assessment is critical

Judging from some of the debates on the use of data in schools, one might think that the only reason for prospective teachers to learn how to interpret student performance data is to interpret the standardized tests that are mandated by states and the federal government.³ The imperative for teachers to be able to use data should not, however, be limited to such tests. Data from assessments of all kinds have always been central to good instructional practice. Effective instruction requires that teachers have a well grounded sense of student proficiency in order to make a daunting number of instructional decisions, such as making snap judgments in the midst of interactions with students, and planning lessons, be they for the next day, the next unit or the entire school year. Setting aside many of the 1,200 or so instructional decisions teachers make on average each day (“Has this student sufficiently mastered the material so that he can work independently until recess?” “What question can I pose that will get the guys in the back row to perk up?”), the number of instructionally *significant* decisions made daily is about 60 (“Which topics should I include on the unit test?”).⁴

Teachers have always needed a firm grasp of their students’ understanding to make appropriate instructional decisions. Many of the first signals attuned teachers pick up on are not the result of any assessment. A teacher needs to be alert to the less tangible indications that students are not grasping new material, from a furrowed brow to “acting out.” That knowledge must be combined with more tangible information gleaned from assessments in the form of spontaneous questioning, homework assignments, quizzes, tests or state-mandated assessments. From informal oral checks to computer-administered tests, assessment provides the basis for instructional decisions, making it critical even in

3 This report does not touch on the need for teachers to understand the use of student performance data in evaluating their own performance, but that understanding is also critical.

4 Hosp, J. L (December 2010). *Linking assessment and instruction: Teacher preparation and professional development. A TQ connection issue paper on improving student outcomes in general and special education.* Washington, DC: National Comprehensive Center for Teacher Quality. (Access at: http://www.tqsource.org/pdfs/TQ_IssuePaper_AssessInstruct.pdf)

the absence of “high stakes” standardized tests.⁵ Further, the use of frequent assessments is a hallmark of the teaching profession in countries whose students outperform our own.⁶ For example, Finland — cited often for a variety of educational features, among them a dearth of standardized testing — trains teachers to use classroom assessment with a high level of sophistication.⁷ Moreover, several recent studies suggest that assessment itself appears to enhance learning through strengthened memory representation.⁸

The new focus on “*data driven instruction*”

While any teacher who casually peruses quiz results and then decides to re-teach a concept could be described as “using data to drive instruction,” that phrase has taken on new meaning in this era of increased school accountability. “Data driven instruction” represents a more organized and collaborative commitment to use data from a variety of assessments — as well as information on student attendance, student engagement, demographics, attendance and school climate — in order to develop or adjust instruction.

Districts may vary in their organizational commitment to data driven instruction, but it is clear that teachers increasingly find themselves not just working in isolation to divine the instructional implications of assessment results, but also working collaboratively with colleagues to use results to improve the performance of individuals, classes, grades or the entire school. For instance, the state of Delaware recently threatened to withhold \$2.5 million from a school district because the district had not properly implemented a 90-minute common planning time in which teachers “take time to strategize and learn to use the voluminous data produced by the state’s new computer assessments to help students in areas where they are struggling most.”⁹

5 The stakes of high stakes assessment were recently raised by research connecting high standardized tests scores with a wide variety of measurable positive outcomes in adult life: Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood* (Working Paper 17699). Cambridge, MA: National Bureau of Economic Research. (Access at: <http://www.nber.org/papers/w17699>)

6 Tucker, M. S. (May 24, 2011). *Standing on the shoulders of giants: An American agenda for education reform*. Washington, DC: National Center on Education and the Economy. (Access at: <http://www.ncee.org/wp-content/uploads/2011/05/Standing-on-the-Shoulders-of-Giants-An-American-Agenda-for-Education-Reform.pdf>)

7 Sahlberg, P. (2011). *Finnish lessons: What can the world learn from educational change in Finland?* New York: Teachers College Press. See p. 66 for a discussion of assessment.

8 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., & Rodeiger 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). *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 36(1), 233-239.

9 Dobo, N. (December 16, 2011). *Delaware tells Brandywine schools to make time for teachers to plan*. Delawareonline.com (Access at: <http://www.delawareonline.com/article/20111216/NEWS/112160335/Delaware-tells-Brandywine-schools-make-time-teachers-plan>)

“Data driven instruction” represents a more organized and collaborative commitment to use data from a variety of assessments and other sources in order to develop or adjust instruction.

The school districts in the nation that do best in the face of the challenge of educating disadvantaged students have become obsessive about using data to drive instruction. Two of the last three winners of the highly coveted Broad Prize for Urban Education, an award to honor “urban school districts that demonstrate the greatest overall performance and improvement in student achievement while reducing achievement gaps among low-income and minority students,” each demonstrated a relentless focus on data.¹⁰

The 2011 Broad Prize winner, **Charlotte-Mecklenburg Schools (CMS)** in North Carolina, narrowed achievement gaps between African-American and white students in reading and math at all school levels. The district’s intense approach to data-based decision making is one factor in this success: All educators in CMS are trained in Data Wise, a structured system to improve instructional and organizational practices based on data, and every school has a data team. Teachers, administrators and coaches can access a wide variety of learning data on the district’s online portal, which gives them the capacity to meaningfully track student performance and adapt quickly to learning difficulties.

Aldine Independent School District (Texas) won the Broad Prize in 2009, with a strong record of academic success for its students and a student performance data initiative that parallels CMS’s. The district developed its own online curriculum and assessment database (in part fed by data from frequent common student assessments), which provides teachers, instructional leaders and district staff with valuable same-day, aggregated and disaggregated student achievement information. This data empowers them to monitor and modify instruction until students demonstrate that they have mastered a subject. District staff can use the data to identify students who need special help, such as students with disabilities or English language learners, and can then use this information to implement targeted staff development in schools where problems are detected. Individual professional development plans are also shaped by test data.

It is fair to say that the school districts in the nation that do best in the face of the challenge of educating disadvantaged students have become obsessive about using data to drive instruction.

Ideally, the training that every aspiring teacher receives would fully prepare them for the growing demands of data driven instruction, employing all potential sources of data that could bear on the academic performance considered, including attendance records, school climate issues such as suspension rates, the nature of needs posed by a variety of English language learners and so on. This report, however, focuses more narrowly on the training teacher candidates receive only in using data derived from *student assessments* — ranging from classwork practice to state tests — to improve instruction. All references to “data” refer to data from such assessments. References to coursework that addresses assessment and assessment data will be referred to as “assessment coursework.”

¹⁰ http://www.broadprize.org/past_winners/2011.html

Previous research

There has been little research on the efficacy of teacher training in assessment. A search of peer-reviewed journals over the last decade produced 26 studies, the vast majority of which do not meet common standards for research. These studies generally have a very small number of subjects (often the researcher’s own students) and make no attempt to examine the effects of any “treatment” on teacher effectiveness. The core finding of these weak studies is that teachers have difficulty analyzing data from classroom assessments and are therefore unable to use the data to guide instruction. The small number that consider the effects of assessment coursework on teacher effectiveness attempt to do so without using data from externally validated data sources. (A list and categorization of these studies is found in Appendix 3.)

The evidence for the connection between using data to drive instruction and student performance (sometimes called “an inquiry cycle”) is emerging, just as the practice of using data is emerging.¹¹ At this early stage, however, and using strict definitions of adequate research, an Institute on Education Sciences (IES) report found “no research that assessed the impact on student achievement of using an inquiry cycle, or individual steps within that cycle, as a framework for data analysis... The panel determined that the level of evidence to support [the recommendation for its use] is *low*.” Notwithstanding this conclusion, however, IES did publish a lengthy “practice guide” on the topic, implicitly recommending that educators implement well-designed inquiry cycles.¹²

Since publication of the IES overview, one study has found that a data driven reform initiative in 500 schools located in seven states led to district-wide improvements in student mathematics achievement, but no such improvements in reading achievement.¹³

11 Black and Wiliam reviewed 250 studies of formative assessment and claim that it has a large potential impact, to arrive at this conclusion, but since these studies are older (they were conducted between 1988 and 1998) when most education research had weaker designs than they do today, an untold number may be suspect. Only the fact that there was a consistent finding for such a very large number of studies lends credence to this conclusion. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-74.

12 National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. (September 2009). *IES practice guide: Using student achievement data to support Instructional Decision Making*. Washington, DC: U.S. Department of Education. (Access at: http://ies.ed.gov/ncee/wwc/pdf/practice_guides/dddm_pg_092909.pdf)

13 Carlson, D., Borman, G., & Robinson, M. (2012). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis*, 33(3), 378-398. A soon-to-be-released study by the Council of the Great City Schools found modest improvements in middle school mathematics and elementary reading.

The evidence for the connection between using data to drive instruction and student performance is emerging, just as the practice of using data is emerging.

A professional imperative from “day one”

U.S. Secretary of Education Arne Duncan made data-driven decision making a national priority, requiring states to improve their data systems and create high-quality assessments to be eligible for a Race to the Top grant. Federal and state governments have devoted increasingly significant resources to the data-driven decision making reform effort. Since 2005, over \$500 million of federal funding has been spent through the Statewide Longitudinal Data System (SLDS) program for development of states’ technology infrastructures.¹⁴

Fortunately, progress has been made. The Data Quality Campaign (DQC) released a report in December 2011 indicating that every state now has the ability to match individual students’ test records from year to year in order to measure academic growth.¹⁵

Districts, too, have been making significant investments. For example, Aldine Independent School District, the 2009 Broad Prize winner, invested \$930,000 over a six-year period in its online curriculum and assessment database — a significant sum given that the district has only 64,000 students.¹⁶

While the data produced by district and state systems may have value in terms of accountability, the full educational potential of these systems can only be realized if teachers have the capacity and motivation to use the data they make available. Yet anecdotal reports indicate that teachers are not taking full advantage of the data that assessments yield.¹⁷ Even among teachers that consider themselves heavy users of student performance data, less than half see data that originate *outside* of their own classroom (such as data from district or state-required tests) as very important. Among teachers who report being light users of data, the proportion of those who see such external data as important is fewer than one in four.¹⁸

14 National Center for Education Statistics. (2010). *Statewide longitudinal data systems grant program*. (Access at: <http://nces.ed.gov/programs/slds/>)

15 Sparks, S. (2011, December 7). Survey Shows Nearly All States Can Track Data on Students. *Education Week*. While we agree that much progress has been made by states, our 2011 *State Teacher Policy Yearbook* (http://www.nctq.org/stpy11/reports/stpy11_california_report.pdf) discusses deficiencies in California’s capacities.

16 According to Ben Wilson, Aldine’s Assistant Superintendent of Community and Governmental Relations, approximately 70 percent of these expenditures were funded by the district, with the remaining 30 percent funded by grants.

17 Of course, the responsibility for poor use of data and data systems is shared between teachers and their districts when the district’s system is poorly designed. For example, New York City spent more than \$80 million on data system design and development, but an audit revealed that less than half of educators accessed the system in spring 2011. (Access at: http://www.comptroller.nyc.gov/bureaus/audit/audits_2012/1-23-12_7111-118A.shtm) An online blog reports that New York City teachers complain that the system offers them too little information, and parents say it’s hard to access. The fact that some teachers and administrators do want data is revealed by the fact that enterprising teachers and schools have created alternative systems for their own use that they are also selling to other public schools. (Access at: <http://gothamschools.org/2010/09/15/frustrated-with-citys-data-system-teachers-build-their-own>)

18 Primary Sources: America’s Teachers on America’s Schools: A Project of *Scholastic* and The Bill and Melinda Gates Foundation (2010). (Access at: http://www.scholastic.com/primarysources/pdfs/Scholastic_Gates_0310.pdf) The most recent report in this series issued in 2012 does not change these conclusions. (Access at http://www.scholastic.com/primarysources/pdfs/Gates2012_full.pdf)

Now, every student comes with instructions.

Introducing STAR Enterprise™—our new generation of reading and math assessments. With Core Progress™ learning progressions, they not only tell teachers how a student is doing, but also help guide teachers to make the best decisions for future engagement. Fast, accurate, and easy to administer, the suite of STAR Enterprise assessments help teachers spend more time teaching and less time testing. Give students the instruction they need.

STAR Enterprise
Take an interactive tour at renlearn.com/lp/18602

A November 2011 advertisement in Education Week. The tools for monitoring and reporting on student performance are an increasingly important feature of K-12 education.

Ironically, many teachers seem to see little use for external assessments even considering the purposes for which they are clearly designed: determining whether students have a deep understanding of content and deciding whether to move them to the next grade. In a recent survey, no more than three in 10 teachers ranked these tests as valuable for such purposes.¹⁹

The Washington Post recently printed some reactions by teachers to a column discussing the growing paperwork burden placed on public school teachers. These responses revealed both the ubiquity of “data crunching” and the overwhelmingly negative sentiments about it:

- *It’s all about data these days, and our professional development days are all about crunching numbers, looking at data and sifting through scores. NEVER is there any talk about how to create engaging and effective and interesting lessons.*
- *Lost in all of this insanity of being driven by the data is the teacher’s cry in the wilderness: Where is the responsibility for the student to learn?*
- *I would argue that in order to improve student achievement, teachers need time to actually instruct children. Instead, too much time has been spent this last quarter on ... assessing students.²⁰*

School district personnel do report that these attitudes can shift with evidence of the positive impact on student learning of data driven instruction.

Identifying a core for assessment knowledge

School districts, states and teacher preparation training programs have yet to establish what knowledge a new teacher should have in order to enter a classroom with some facility for applying data to improve classroom instruction. In fact, the field has struggled to incorporate data-driven decision making into its program sequences. As researchers Ellen Mandinach and Edith Gummer wrote:

The field of data-driven decision making falls between the cracks of traditional courses. It is not introductory statistics or measurement; it is not instruction, pedagogy, or methods. Data-driven decision making lies at the intersection of those courses.

— *The Complexities of Integrating Data-Driven Decision Making into Professional Preparation in Schools of Education: It’s Harder Than You Think*²¹

In 2010, NCATE, the largest accreditor of teacher preparation programs, published a report on assessment in teacher preparation²² with the general recommendation that “candidates

19 Survey results can be accessed at <http://www.nwea.org/every-child-multiple-measures>

20 What’s lost in the paperwork. (2011, November 21). *The Washington Post*, p. B2.

21 Mandinach, E. B., & Gummer, E. S. (2011, May 12). *The complexities of integrating data-driven decision making into professional preparation in schools of education: It’s harder than you think*. Report from an Invitational Meeting Presented to the Spencer Foundation. (Access at: http://educationnorthwest.org/webfm_send/1133)

22 *Assessment as a critical element in clinical experiences for teacher preparation*. (Access at: <http://www.ncate.org/LinkClick.aspx?fileticket=oo50CSYDEFM%3D&tabid=715>)

“In my experience, when teachers are able to see first-hand how the use of data can impact the achievement of *their* students they can be convinced of the value of data driven instruction.”

— Abbey Goldstein
Manager, Strategic Initiatives
District of Columbia
Public Schools

be presented with multiple and rich course material in their preparation that will enable them to become *assessment-literate* and *data-wise*.” It did not however, provide specific suggestions as to how such material could be integrated into existing coursework or provided through new coursework.²³

After a review of the literature and the opinions of experts in the field,²⁴ we delineated the following three domains of knowledge needed by teacher candidates in the area of assessment preparation:

1. How to measure student performance using assessments: “**Assessment Literacy**”
2. How to analyze student performance data from such assessments: “**Analytical Skills**”
3. How to use student performance data from assessments to plan instruction: “**Instructional Decision Making**”

These three domains of knowledge formed the basis of our examination of institutional practice in assessment. As we sifted through the syllabi for professional coursework that 180 programs require of their teacher candidates, we classified a course as relevant if it provided content associated with one or more of the three domains. We then proceeded to rate the comprehensiveness of a program’s coverage.

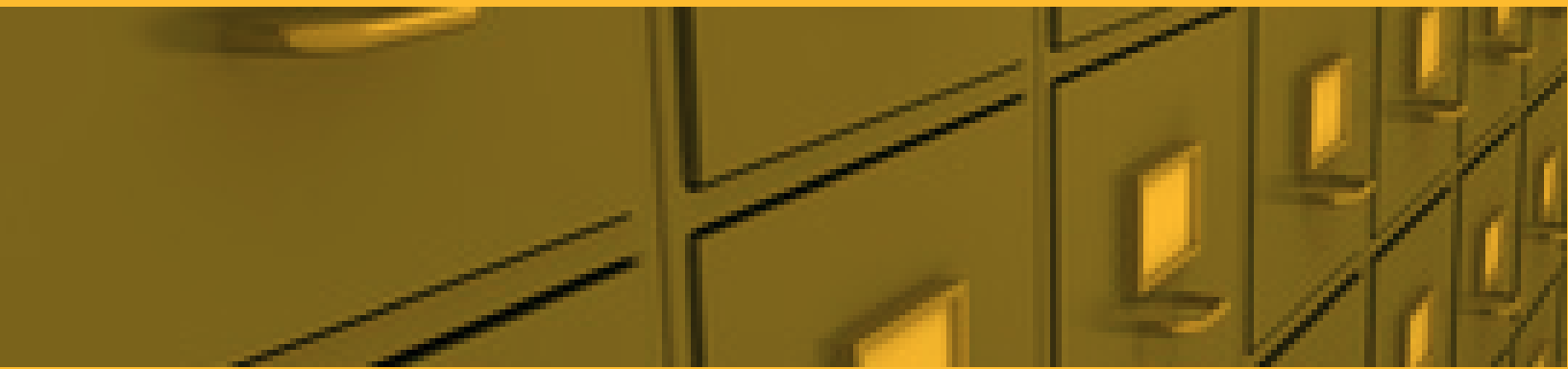
A full discussion of each domain, the rubrics used for evaluation in each of the three domains, and descriptions of the relevant features of sample programs under each rating category are found in Appendix 2.

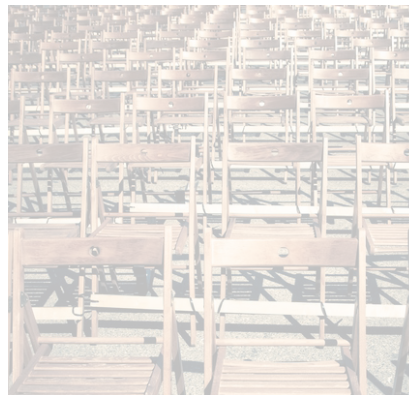
The three domains of assessment knowledge needed by teacher candidates:

1. Assessment Literacy
2. Analytical Skills
3. Instructional Decision Making

²³ Appendix 4 contains a full review of state regulations, institution accreditation standards and professional standards, examining the degree to which they have articulated the requisite teacher knowledge and performance.

²⁴ Notably Dr. John Hosp, Associate Professor, Department of Teaching and Learning, University of Iowa; Dr. Barry Kaufman, President, BK Education Consulting Services; and Dr. Ellen Mandinach, Senior Research Scientist, WestEd.





Methodology

The study sample

The sample for this preliminary report is 180 teacher preparation programs for both elementary and secondary teacher candidates, housed in 98 institutions of higher education (IHE) in 30 states. These institutions were selected for the sample because they were in the first states responding to open records requests from NCTQ to provide syllabi associated with required professional coursework toward initial certification, an aspect of our *National Review of Teacher Preparation Programs*.²⁵

The graphics that follow illustrate the characteristics of the sample relative to the total population of institutions offering teacher preparation programs, as well as illustrating the distribution of the sample relative to institutional selectivity in admissions as measured by average SAT scores of all students.

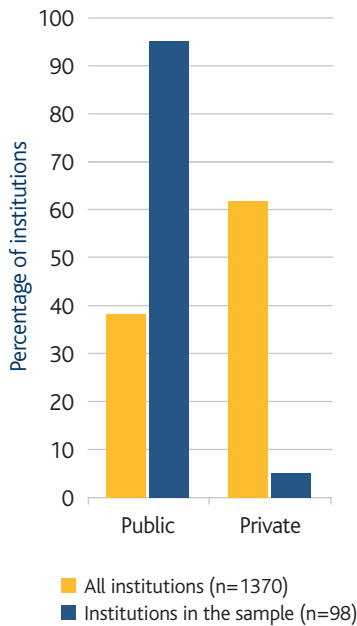
While the sample is overwhelmingly composed of public institutions, we have no reason to believe that this biases results. The results from our analysis of assessment coursework in Illinois²⁶ suggest that there is little difference in the practices of public and private institutions.²⁷

25 <http://www.nctq.org/p/edschools/home.jsp>

26 Greenberg, J., & Walsh, K. (2010). *Ed school essentials: A review of Illinois teacher preparation*. Washington, DC: National Council on Teacher Quality. (Access at: <http://www.nctq.org/edschoolreports/Illinois>)

27 While the rating process in Illinois was not exactly the same as used here or as will be used in the *National Review*, programs in private institutions in Illinois had average ratings of 3.0 on a 4-point scale, whereas ratings of public institutions averaged 2.8.

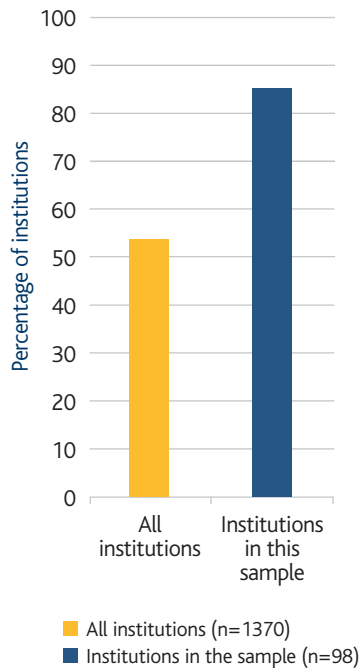
Public and private institutions*



*Data from Title II reports

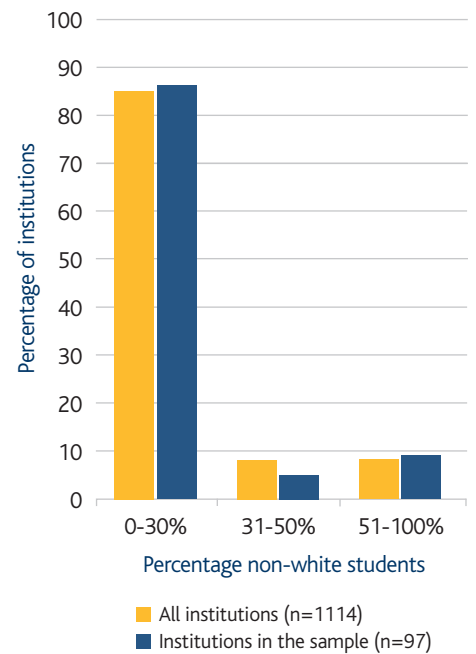
Data were obtained mostly through open records requests to which only public institutions must respond. As a result, while just over one-third of the IHE's offering teacher preparation are public institutions, about 95 percent of the IHEs in the sample are public.

National accreditation of institutions



Because many states require that IHEs offering teacher preparation seek accreditation, a disproportionate number of accredited IHEs are public. This may explain why the proportion of accredited IHEs in the sample is much larger than in the proportion of accredited IHEs overall.

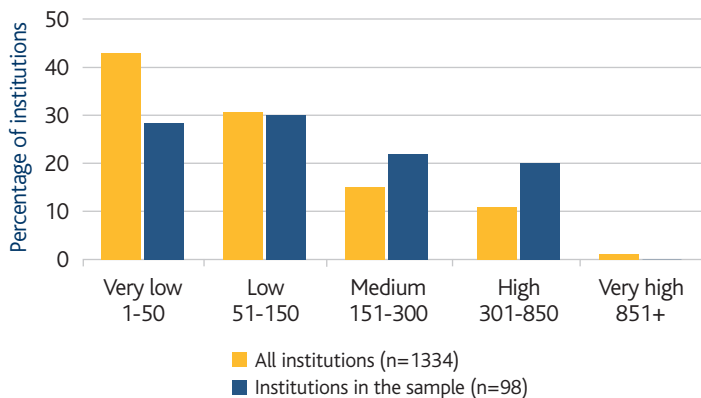
Diversity of institutions*



*Data from Title II reports

In terms of minority enrollment, the sample is representative of the population of IHEs offering teacher preparation.

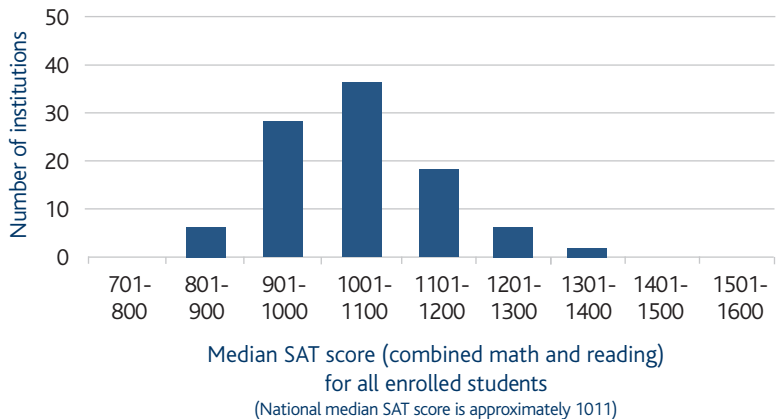
Number of teachers produced annually by institutions*



*Data from Title II reports

The distribution of IHEs offering teacher preparation is dominated by institutions producing 50 or fewer graduates annually in all programs, many of them private. The fact that the sample contains a disproportionate share of public IHEs may explain why the sample has a more even distribution across four of five categories of production.

Admissions selectivity of institutions in the sample*



* Data obtained from IPEDS for all but four IHEs for which scores were obtained from Cappex.

Selectivity information based on average student SAT scores is not available for a large number of IHEs. This only illustrates the distribution of SAT scores for the sample.

The methodology for selection of coursework

Any required course was considered relevant for further analysis if the title of the course or its description in the IHE's catalog indicated that the topic of assessment might be a significant feature of instruction and/or practice.²⁸ The syllabi for all such courses were reviewed.

Some courses were not classified as relevant to the aims of this study even though their title and/or description referenced the topic of assessment. These include courses in educational psychology that are taught in a psychology department (rather than an education department) and that are not designed for teacher audiences. These courses' treatment of assessment would not be sufficiently contextualized to K-12 schooling to include in this study. Courses in which literacy was the main topic were also excluded.²⁹

For the purpose of evaluating programs preparing elementary and secondary teachers, all professional courses referencing assessment in the title or course description (with the exceptions noted above in educational psychology and literacy) were evaluated. For the purpose of evaluating programs preparing secondary teachers, one of four pathways for secondary certification was randomly selected (English, mathematics, science or social studies) and the methods course for the path chosen was also evaluated, regardless of whether assessment was mentioned in its title or description.³⁰

After deciding which courses appeared to be relevant, programs were evaluated on the basis of the following information:

- The course syllabi stating the course objectives, lectures, assignments, textbooks and required readings;³¹
- Information provided by publishers' descriptions of the content of required textbooks; and
- Requirements for capstone projects (often called "teacher work samples" or "portfolios").³²

28 We had anticipated that institutions would assist in the identification of assessment coursework through a voluntary curriculum mapping process, but the lack of cooperation from institutions in the *National Review* sample made this impossible. Institutions' requests for excessive payments for data also made it impossible to obtain all professional coursework syllabi for all schools to examine them for coverage of assessment. It is possible that courses treat assessment but the title or course description did not so indicate. To ascertain the potential for the coverage of assessment that is not revealed by title or course description, we conducted an analysis of coursework in 18 programs in the sample for which we had all syllabi available for evaluation. Our analysis indicates that titles and descriptions of many methods courses do not mention assessment, even though assessment is the topic of one or two lectures. However, with the exception of one course in one program, there were no practice assignments on assessment other than in the course of writing lesson plans, and there was no instruction or practice that would be evaluated in the second and third assessment domains that will shortly be described. Both because the program in which non-lesson planning-related practice was found already had adequate practice and because instruction and practice was related only to the first domain, had this instruction and practice been included in our evaluation, none of the additional coverage of assessment in these courses in the 18 programs would have changed the overall rating of any of the programs in which it was found.

29 Assessment in literacy instruction (ranging from typical progress monitoring to the assessment that is integral to systematic programs such as "Response to Intervention") requires a more specialized evaluation than provided here. Our evaluation of literacy coursework against our early reading standard addresses assessment and assessment strategies. For example, we evaluate whether instruction on assessment strategies focuses on systematic skill building as opposed to approaches such as running records.

30 It is reasonable to assume that assessment-informed planning would be addressed in the separate subject-specific methods course for each pathway to certification. In other words, it is valid to evaluate only a sample of methods courses in a program in order to produce a general evaluation of methods coursework because it parallels the experiences of teacher candidates.

31 Several researchers evaluating teacher preparation in literacy assessment-evaluated programs have used this same source of data: McCombes-Tolis, J., & Spear-Swerling, L. (2011). The preparation of pre-service elementary educators in understanding and applying the terms, concepts, and practices associated with response to intervention in early reading contexts. *Journal of School Leadership*, 21(3), 360-389.

32 Teacher Performance Assessments (TPAs) were also evaluated as capstone projects. For all projects, we accounted for any assessment-related project assignments, including whether these projects required that candidates submit artifacts from earlier coursework that reflected assessment assignments or entailed new assessment assignments.

On this basis, a total of 455 courses were evaluated, anywhere from one to six relevant courses for each program. The average number of courses in each program that reference assessment is 2.5.

The validity of using syllabi

Analyses of syllabi have long been an accepted part of the evaluation of teacher preparation by state agencies, accrediting organizations and multiple research studies. The NCTQ methodology mirrors this approach, looking only for topics that are considered so essential that an instructor would not have failed to mention them on the syllabus if she or he intended to teach the topics at all.³³ Even in cases when syllabi are cryptic, careful examination yields valuable information about what teacher candidates are expected to learn and practice. For example, if a syllabus was unclear whether an assignment for an “assessment analysis” included analysis of both classroom and standardized assessments, it was assumed that both were addressed, provided that a lecture delivered before rather than after the assignment was due dealt with standardized tests. In the absence of clear lecture topics or practice assignments, course objectives were used to assess course content.³⁴

A total of 455 courses in 180 programs were evaluated for their coverage of assessment.

33 Teacher educators often object to the use of syllabi to assess the core of instruction. Most objections seem to be based on the proposition that the most important concepts in a course can be so deeply embedded that they are invisible. In the words of one education school dean, if something is a “natural part” of a course it “need not be explicitly identified.” In fact, it is exactly the “natural part” of a course that is routinely expected by students, instructors, college administrators, trustees and accreditors to be described by syllabi that are the topics salient to this analysis.

34 See “Sample program earning a ‘4’” on p. 37 for an example of coursework evaluated in the absence of clear lecture topics.



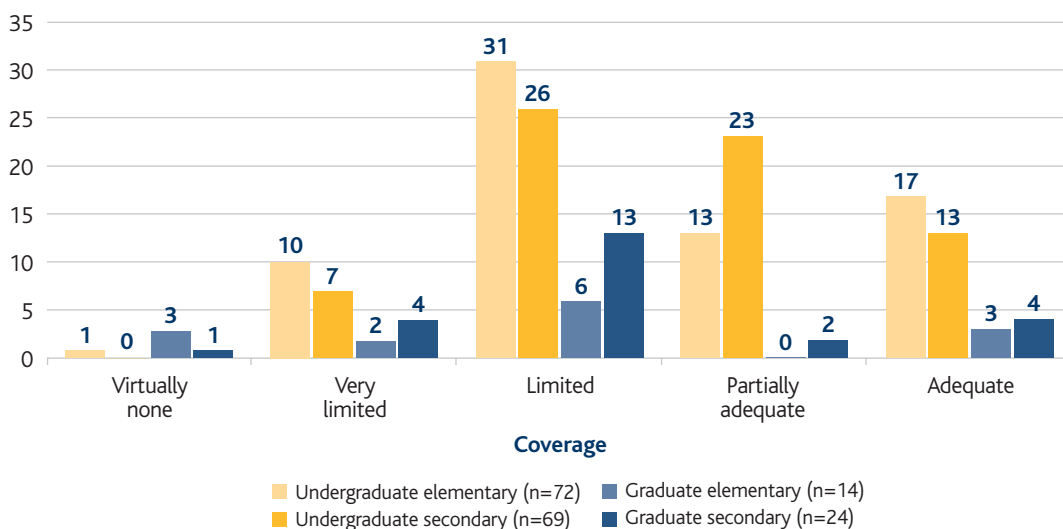
Findings

How adequately does teacher preparation program coursework address Domain #1, Assessment Literacy?

Assessment Literacy

An understanding of the taxonomy of assessment and a familiarity with the different types of classrooms and standardized assessment. For example, teacher candidates should understand the distinction between formative and summative assessment and have developed several types of each form of assessment. Appendix 2 contains a more detailed description of this domain, the rubric used for evaluating programs in this domain area, and descriptions of the relevant features of sample programs in each of the rubric's rating categories.

How adequately does coursework address "Assessment Literacy"?



Only 21 percent of the programs in the sample cover literacy topics adequately, with an additional 21 percent doing so with partial adequacy. More than half of all programs have no, very limited or limited coverage.

In the first domain, only five programs (3 percent) in our sample earned the lowest rating, and one in five programs (21 percent) earned the highest. Findings in this domain are clearly more positive than findings in the two other domains.

Programs that earned the highest score provided:

1. Comprehensive coverage of both classroom and standardized assessment (including key concepts such as validity and reliability), and
2. Practice on developing and scoring assessments.

Most of the remaining programs provide teacher candidates with at least basic exposure to both formative and summative assessment.³⁵ Moreover, in about two-thirds of the programs (67 percent), the coursework contained at least some reference — albeit, in most cases, fairly minor references — to standardized testing. We also noted little explicit skepticism about state assessments and accountability systems, finding it in only six programs (3 percent). Given that a recent survey of teacher educators found that only 24 percent of teacher educators believe it “absolutely essential” to produce “teachers who understand how to work with the state’s standards, tests and accountability systems,”³⁶ this absence of explicit skepticism is commendable.

How adequately does teacher preparation program coursework address Domain #2, *Analytical Skills*?

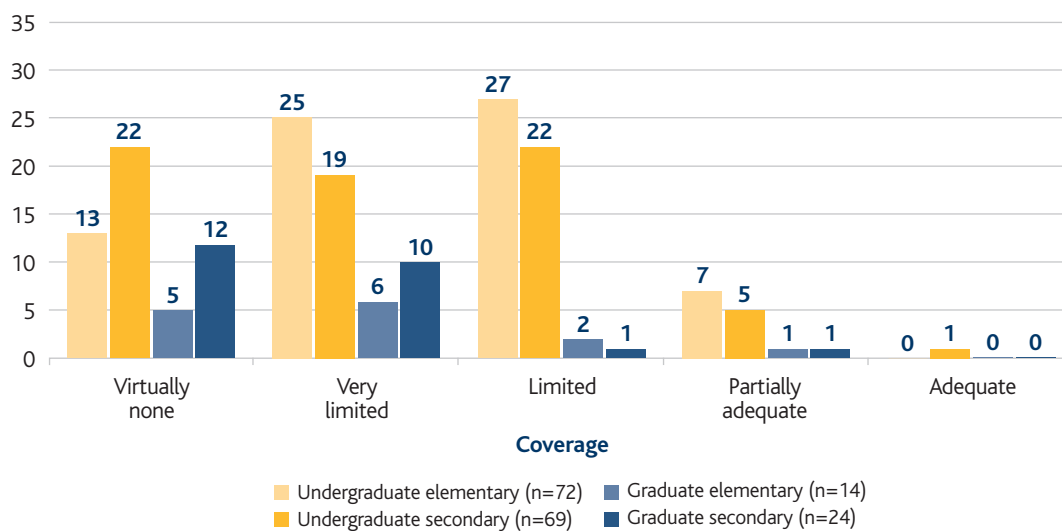
Analytical Skills

Understanding how to dissect, describe and display the data that emerges from assessments. For example, teacher candidates might learn how to disaggregate class results on a series of assessments in a unit to review and present differential subgroup performance in a table or chart. Appendix 2 contains a more detailed description of this domain, the rubric used for evaluating programs in this domain area, and descriptions of the relevant features of sample programs in each of the rubric’s rating categories.

35 Our findings run contrary to the results of a recent ETS study that found formative assessment mentioned explicitly in only three of the 22 programs (14 percent) that ETS examined. (Access at: http://blogs.edweek.org/edweek/teacherbeat/2011/12/how_do_we_train_teachers_to_us.html) In an examination of the first 27 institutions (48 programs) evaluated, (approximately one-quarter of the total programs evaluated), we found explicit references in roughly half of the programs (14/27 or 52 percent). We found implicit references to assessment (e.g., a lecture on “how to determine what students know”) and/or capstone projects that required teachers to design formative assessments in all but one of the remaining 13 programs.

36 *Cracks in the ivory tower?: The views of education professors circa 2010.* (Access at: <http://educationnext.org/cracks-in-the-ivory-tower-the-views-of-education-professors-circa-2010/>)

How adequately does program coursework address “Analytical Skills”?



Less than 1 percent of the programs in the sample cover Analytical Skills adequately, with an additional 8 percent doing so with partial adequacy. The vast majority of programs (92 percent) have no, very limited or limited coverage.

The ratings are considerably poorer in this second domain. Over one-quarter of all of the programs (29 percent) and almost half of the graduate programs (45 percent) were completely deficient, and only one program was found adequate.

The program deemed “adequate” provided:

1. Instruction on analysis of data from both classroom and standardized assessments, and
2. Practice that includes both field-based practice in teams and presentation of results in both quantitative and graphic displays.

In general, coursework appears to provide only the most basic tools for analysis of assessment data and then primarily from classroom assessment. Only 71 programs (39 percent) address standardized testing in the context of data analysis, and fewer — 22 programs (12 percent) — have classwork or homework practice that exposes teacher candidates to the process of analyzing data from standardized assessment. In 72 programs (40 percent), teacher candidates have to demonstrate analytical proficiency through independent “capstone projects” in which they analyze classroom assessment data, but this exercise appears to be candidates’ only practice in 43 programs (60 percent of those programs with such projects). We found evidence of *collaborative* analysis in in-class or capstone project assignments in a very small fraction (4 percent) of the programs.

How adequately does teacher preparation program coursework address Domain #3, *Instructional Decision Making*?

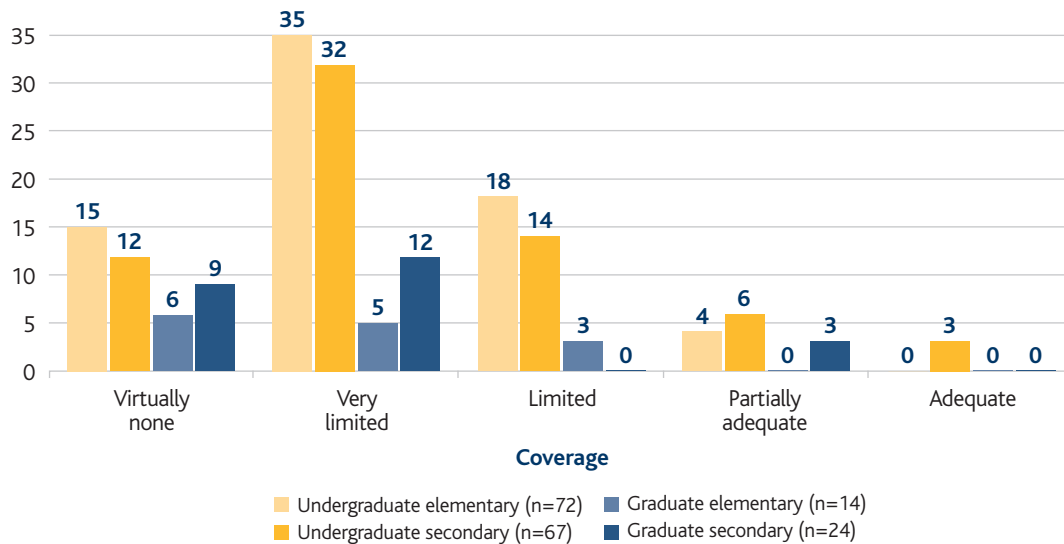
Instructional Decision Making

An understanding of how to derive instructional guidance from assessment data. For example, a teacher candidate preparing to teach secondary mathematics might examine a variety of test results obtained in field work to determine the reasons students have performed poorly on a math problem on the subject of rates. The teacher candidate will need to determine if students lacked a basic understanding of rates themselves or if they were confused about the mixed-number multiplication necessary to compute the rate.* The next steps would depend on the outcome of this examination. They might lead to formative assessment to “zero in” on students’ understanding either of rates or mixed-number multiplication (or both), appropriately focused instruction, and a summative assessment for final student evaluation.

Appendix 2 contains a more detailed description of this domain, the rubric used for evaluating programs in this domain area and descriptions of the relevant features of sample programs in each of the rubric’s rating categories.

* See *Driven by Data: A Practical Guide to Improve Instruction* by Paul Bambrick-Santoyo (John Wiley & Sons, Inc., 2010) for more examples of the triangulation approach to Instructional Decision Making.

How adequately does coursework address “Instructional Decision Making”?



Fewer than 2 percent of programs in the sample cover Instructional Decision Making adequately, with 7 percent doing so with partial adequacy. The vast majority of programs (91 percent) have no, very limited or limited coverage.

In this third domain of Instructional Decision Making, findings remain almost as troubling as in the second domain. Only three programs (all undergraduate secondary) earned the highest score. In only a slight improvement over results in the second domain, almost one-quarter (23 percent) of all programs and over one-third (39 percent) of graduate programs were completely deficient.

Programs found adequate provided instruction and practice on using assessment data to drive instruction that is extensive and includes coursework that is subject-specific.³⁷

Many programs do provide at least cursory exposure to the concept of “formative assessment,” which constitutes a toehold on how to use assessment data for planning.³⁸ Programs generally, however, do not appear to use their methods courses as the vehicles for more nuanced, subject-specific practice. In 42 elementary programs (48 percent of such programs), using assessment to drive instruction appeared to be virtually absent from methods courses, meaning that no expert taught teachers how to examine assessment data with the benefit of subject matter-specific insight.³⁹ For example, a science methods course and a mathematics methods course could provide different but equally valuable insights into subtle misconceptions and misinterpretations that might underlie incorrect answers, or how assessment results might suggest necessary instructional scaffolding.

In 54 secondary programs (58 percent of such programs), there was either no subject-specific methods course, or there was a subject-specific methods course that did not address assessment. Again, this finding means that a teacher candidate in a social studies certification route, for example, may not have the advantage of an instructor with content expertise as she tries to evaluate the content validity of a test she has developed with original multiple choice questions. Even in programs in which assessment was a topic of methods coursework, it was generally a very minor topic at best. For example, only one lecture might address assessment in the subject at hand; teacher candidates might be required to complete an assignment entailing only one assessment, with no explicit connection to how its use might inform instruction.

Lastly, we found evidence of *collaborative* Instructional Decision Making practice in in-class or capstone project assignments in a very small fraction (8 percent) of the programs.

The bar to earn a passing rating in this study was set low.

How adequately does teacher preparation coursework address all three domains of assessment?

Before discussing the overall performance of programs, it is important to note that the bar to earn a passing rating in this study was set low.

Just how low? One relevant objective or lecture was sufficient to categorize a program as addressing standardized assessment. Several assignments were sufficient to categorize a program as having “adequate practice” in a particular area. Some might reasonably suggest that teachers, especially in this era of accountability and assessment, need far more exposure and practice than this standard suggests—a fair criticism in our judgment.

37 This standard pertains to secondary programs. The analogous standard for elementary programs requires that instruction and practice are evident in all four core elementary subjects: language arts, mathematics, science and social studies.

38 See Appendix 5 for a tutorial on assessment taxonomy.

39 As mentioned earlier, we assume that virtually all methods courses address assessment to the minimal extent of requiring that lesson plans include at least summative assessments.

Why do we set such a low bar? We do so to give institutions the benefit of the doubt. Though our examination of course materials is comprehensive, our low bar compensates for our inability to audit the actual course and the risk we face that a course might actually cover material that was not spelled out in the syllabus. This approach means that some programs earn higher marks than perhaps they should have. But it also means that our margin of error is so substantial that there should be little doubt that a program designated as inadequate is in fact inadequate.

Our overall conclusion is that while assessment is addressed to some extent in all but five of the 180 programs we examined, only six programs (3 percent) provide preparation that can be deemed adequate: four elementary programs and two secondary programs.

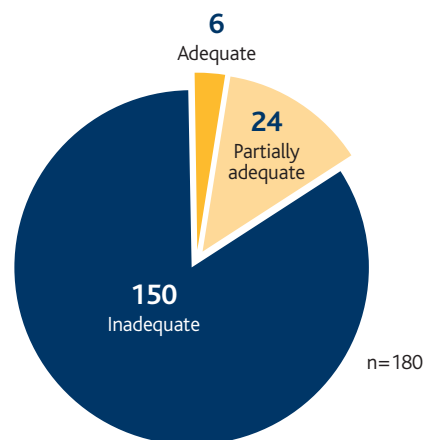
Only six programs (3 percent) provide preparation that can be deemed adequate.

When it comes to assessment, what is an “adequate” program?

- 1) The program provides a foundation in Assessment Literacy, with attention to the instructional role of standardized tests, particularly the program state’s standardized tests;
- 2) The program requires teacher candidates to prepare formative and summative classroom assessments; and
- 3) The program builds a relatively strong foundation for interpreting and applying data, *both individually and in teams*, from standardized and classroom assessments in order to inform and adjust instruction.

These three parts are based roughly on the three domains, but combining them into an overall metric moves us away from a simple aggregation of ratings in each domain. Frankly, this allows us to note a program as “adequate” that may not have received the highest rating on any of the domains but still provides a “*relatively strong*” foundation in all.

How adequate overall is teacher preparation program coverage of assessment?



Only 3 percent of programs in the sample cover assessment adequately, with 13 percent doing so with partial adequacy. Over three-quarters of programs (83 percent) have inadequate coverage.

The most common weaknesses:

- As noted, programs were rated much more poorly on the second (Analytical Skills) and third (Instructional Decision Making) domains than on the first (Assessment Literacy). Even among the programs earning the highest rating on the first domain, few were adequate on the domains of Analytical Skills and Instructional Decision Making.
- While about two-thirds of programs address standardized assessments, coverage is found largely in the first domain of Assessment Literacy. Even programs that address standardized assessments rarely discuss data from standardized tests in the context of Analytical Skills or Instructional Decision Making, or have relevant practice.
- Programs rarely mention the use of data by teachers for *school* improvement purposes, (as opposed to improvement of the performance of their own students).
- There is little evidence of collaborative practice being modelled or promoted.

We found that programs in the same institution may be internally *inconsistent*:

In only one of the three institutions in which we evaluated all four program types were all the programs rated the same—in this case, unfortunately, as “inadequate.”

The impact of state regulation.

In many areas of teacher preparation, NCTQ has found that state regulations governing the content of teacher preparation have little impact in actual practice.⁴⁰ In the policy brief we released earlier this year, our analysis of a smaller sample of programs suggested that state regulation on preparation for assessment might be an exception to this general rule. With the full sample in hand, however, we can once again confirm that state regulations do not affect what teachers learn in their preparation programs.⁴¹ As noted in our recommendations, this is likely due to the fact that states do not require candidates to pass tests in this subject to become licensed.

Even among the programs earning the highest rating on the first domain, few rated well on the domains of Analytical Skills and Instructional Decision Making.

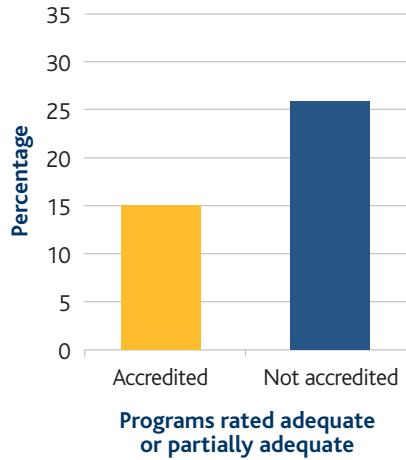
⁴⁰ See Appendix 4 for a discussion of relevant state regulations.

⁴¹ The 180 programs evaluated to date are found in 30 states, 18 of which have regulations that at least address to some extent the responsibility of teacher preparation programs to address assessment. (See graphic in Appendix 4.) In all states, whether regulations are relatively strong or weak, ratings on programs within the state are mixed in states for which more than one program was evaluated. However, only a slightly higher proportion of programs are “adequate” or “partially adequate” (18 percent) in states that are relatively stronger on regulation than in states that are not (15 percent).

The impact of accreditation.

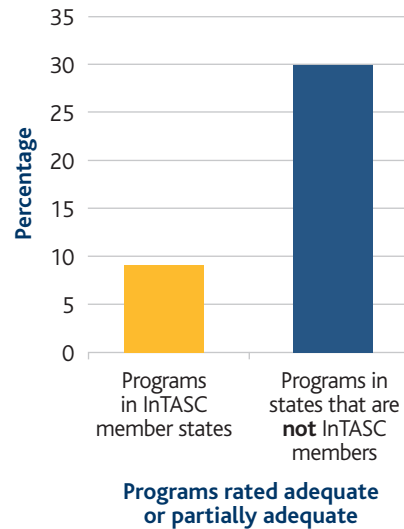
In previous analyses of various areas of teacher preparation we have not found a significant difference between accredited and non-accredited institutions. In this area, we have, and the difference points to a negative impact of accreditation for the 157 programs that are accredited by either NCATE or TEAC out of this sample of 180 programs.⁴²

Does accreditation make a difference?



Twenty-six percent of the programs in non-accredited institutions are rated as “adequate” or “partially adequate,” whereas only 15 percent of the programs in institutions accredited by NCATE or TEAC fall into either of these two rating categories.

Does a state’s adoption of InTASC professional standards make a difference?



Only 9 percent of the programs in states that are members of InTASC are rated “adequate” or “partially adequate.” A much higher proportion (30 percent) of programs in states that are not members fall into either of these two rating categories.

The impact of professional standards

As with accreditation, if there is a difference, it appears to indicate a *negative* impact of professional standards for the 109 programs in the 21 states in the sample which are members of InTASC and subscribe to its professional standards.⁴³

42 See Appendix 4 for a discussion of relevant accreditation standards.

43 See Appendix 4 for a discussion of relevant professional standards.



Recommendations

We fully recognize that assessment and use of assessment data is such a large and important topic for teacher candidates that no matter how adequate a program may be in laying a foundation, continued training will be necessary to build and hone knowledge and skills during induction and throughout classroom teaching. But the need for in-service training does not negate the need for foundational knowledge to be also laid during pre-service. Numerous levers are available to provide both increased pressure on and increased support for teacher preparation programs to lay a better foundation for continuing professional development in the area of assessment.

Federal Government

Provide more legislative guidance

The federal government should encourage better teacher candidate preparation in data driven instruction through:

1. Amending Title II of the Higher Education Act (HEA), providing suitable incentives for teacher preparation programs to prepare teacher candidates to understand and use assessments and assessment data;
2. Requiring that a portion of State Longitudinal Data System (SLDS) funds be used for training on assessment and data skills; and
3. Allocating a percentage of funds under Title II of the Elementary and Secondary Education Act (ESEA) for training.

Amendments to HEA and ESEA, respectively, could also

1. Provide states with incentives to ensure that their program approval authority reinforces the importance of teacher data skills; and
2. Encourage innovative state and local practices in data driven instruction.

Invest in research

There are a number of issues related to data driven instruction that warrant exploration. Federal policy should support research into its effects on student achievement, the knowledge and skills teachers need to become expert, the conditions within schools necessary for it to have maximum impact, the collection and dissemination of best practices related to pre-service and in-service training, and the compilation and meta-analysis of existing research on assessment, data driven instruction and related topics.

We present two different and potentially complementary approaches that might increase the accountability of preparation programs, one for states and one for school districts:

States

Increase accountability of preparation programs

Teacher licensing tests are perhaps the most effective and efficient means available to drive significant change in teacher preparation programs. One possible reason that assessment is addressed so poorly in teacher preparation programs is that it is not covered at any depth in the licensing tests required by most states. As a first step to remedying this situation, the **Council of Chief State School Officers (CCSSO)** could evaluate the scope of topics addressed in those licensing tests that do include assessment to make recommendations to states about how licensing tests might be improved in this area. If warranted, CCSSO could sponsor a consortium of states who wish to aggressively promote licensing test enhancements.

School Districts

Ensure a better applicant pool for hiring

Of all the stakeholders in the country’s educational system, **school districts** have the greatest vested interest in hiring teachers with a solid foundation in assessment. Now – of necessity – districts must hire new teachers who no sooner arrive at their new schools than they need immediate professional development on a variety of assessment tools and tasks necessary for school improvement. Districts hiring a large share of graduates from any given teacher preparation program might use their bargaining power to bring the program to the table to ensure that its teacher candidates arrive with a better foundation in all assessment domains.

Absent a state effort to focus pedagogical licensing tests on the topic of assessment knowledge that teachers need, we recommend that districts administer their own test, one that applicants must take as a condition of hiring. Such a test might be designed under the supervision of, for example, the **Council of the Great City Schools**. It may be the most efficient step toward persuading states that assessment knowledge needs to be an integral part of any licensing test.

Foundations

Develop instructional data sets and model curricula

A variety of graduate programs for teachers and administrators offer coursework in the data driven instruction model that can be taken in degree programs or as stand-alone in-service training.⁴⁴ Participants must bring to these programs their own district’s student performance data inventory. Course assignments have participants working directly with their own data. Teacher educators in *initial* certification programs seeking to improve their courses, however,

⁴⁴ As mentioned earlier, all teachers in the Charlotte-Mecklenburg school district – a Broad Prize winner – have taken one of the most well-known courses, Harvard’s “Data Wise” course.

Now — of necessity — districts must hire new teachers who no sooner arrive at their new schools than they need professional development on a variety of assessment tools and tasks necessary for school improvement.

would have a hard time replicating these inventories. Neither the instructor nor enrolled teacher candidates has access to the appropriate data, and fabricating data is not easy.⁴⁵

Innovation-minded teacher educators would welcome synthetic but realistic elementary and secondary school data sets that contain student performance data from a typical range of classroom, district and state assessments. These data sets could form the basis of a very rich collection of simulations and exercises designed to develop expertise in Assessment Literacy, Analytical Skills and Instructional Decision Making in existing coursework or in model course curricula developed by foundations and their partners.⁴⁶

School data sets that contain student performance data from a typical range of classroom, district and state assessments could form the basis of simulations and exercises in teacher preparation coursework.

How might student performance data be used in instruction?

Consider as a hypothetical a case example in an Institute for Education Sciences practice guide on “Using State Achievement Data to Support Instructional Decision Making.”* The case example describes how a group of elementary teachers might analyze data on student performance in mathematics: The teachers find that students are weakest on state tests in geometry skills related to shapes and measurements, but data on district interim assessments do not reveal the source of the problem. It is only when student workbooks are consulted that the teachers surmise that the students’ difficulties relate to the need to combine shapes for some calculations. This leads the teachers to generate some relevant classroom lessons. The next interim district test shows improvement on the type of problems that the teachers had analyzed collaboratively.

Teams of teacher candidates could simulate this same exercise in coursework. They could be provided a data set containing 1) fabricated but realistic 4th grade student test scores on the mathematics section of a state test and interim district assessments taken throughout the year, 2) appropriate statistical descriptions and graphic displays, and 3) examples of student work designed to illuminate possible student failings and misconceptions. Working together, the teacher candidates could develop appropriate lesson and assessment plans, just as they might soon be doing with colleagues when they begin teaching.

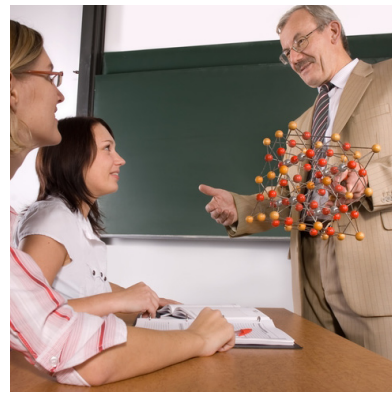
Developing data sets for the core subjects at both the elementary and secondary levels could provide ample practice for teacher preparation coursework to be used ideally in courses in which faculty with assessment expertise and faculty with pedagogical expertise could coordinate instruction or team-teach.

* National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. (September 2009). *IES practice guide: Using student achievement data to support Instructional Decision Making* (p. 17). Washington, DC: U.S. Department of Education.

⁴⁵ Several instructors of graduate level assessment coursework informed us that they have attempted to develop such a data inventory for course participants who had no access to data and that they had not found the results satisfactory because the task of replicating the massive amount of data now available to many school administrators and teachers is so difficult.

⁴⁶ The DQC reports that discussions are underway between the New Hampshire Department of Education and teacher preparation programs in the state to develop a college course around effective use of data to plan instruction. (Access at: http://www.dataqualitycampaign.org/resources/field_profiles/NH.A9)





Conclusion

We find evidence of a serious misalignment between what school districts seek in the skills of their teacher corps and the training provided by teacher preparation programs. It is noteworthy that only two of the 455 assessment courses we identified appear to exploit local school district expertise, in this case by having a representative from a local district lecture, presumably on student performance data in the district context.

In too many programs, assessment coursework centers only on the tests that teachers have always administered, preparing teacher candidates to develop and use assessment data to improve their students' performance in an insular environment. As important as this type of preparation continues to be, it shortchanges teacher candidates because it does not represent the environment of schools in this century. New teachers undoubtedly find themselves confronting data presentations using unfamiliar terms and concepts. The flow of data they currently have to deal with will only grow larger in 2014, as new assessments tied to the Common Core state standards are first administered. Today's schools demand teachers who can comfortably understand and utilize — both individually and collaboratively — a full range of classroom and standardized data, whether the data relate to their own students or to all the students in their school. Preparing them for anything less is unfair to teacher candidates, as well as to the many students they plan to teach.

**This report is available online at
www.nctq.org/p/publications/docs/assessment_report.pdf**



National Council on Teacher Quality

1420 New York Avenue, Suite 800
Washington, D.C. 20005
Tel: 202 393-0020 Fax: 202 393-0095
Web: www.nctq.org

The National Council on Teacher Quality advocates for reforms in a broad range of teacher policies at the federal, state and local levels in order to increase the number of effective teachers.

Subscribe to NCTQ's free monthly electronic newsletter, *Teacher Quality Bulletin*, (www.nctq.org/p/tqb/subscribe.jsp), visit our *Pretty Darn Quick* blog at www.nctq.org/p/tqb/pdq.jsp or follow us on Facebook or Twitter to stay abreast of trends in federal, state and local teacher policies and the events that help to shape them.