Reviewers:
Marilyn Cochran-Smith, Boston College
Elizabeth Stringer Keefe, Lesley University
Wen-Chia Chang, Boston College
Molly Cummings Carney, Boston College

June 2018

National Education Policy Center
School of Education, University of Colorado Boulder
Boulder, CO 80309-0249
(802) 383-0058
eipc.colorado.edu
Acknowledgements

**NEPC Staff**

Kevin Welner  
Project Director

William Mathis  
Managing Director

Alex Molnar  
Publications Director

---


**Funding:** This review was made possible in part by funding from the Great Lakes Center for Educational Research and Practice.

This publication is provided free of cost to NEPC’s readers, who may make non-commercial use of it as long as NEPC and its author(s) are credited as the source. For inquiries about commercial use, please contact NEPC at nepc@colorado.edu.

The authors of this review are members of PROJECT TEER (Teacher Education and Education Reform), which is a group of teacher education scholars and practitioners who have been studying U.S. teacher education in the context of larger reform movements since 2014. Led by Marilyn Cochran-Smith, the group’s work includes several NEPC publications, multiple journal articles, and the book, *Reclaiming Accountability in Teacher Education* (Cochran-Smith, Carney, Keefe, Burton, Chang, Fernández, Miller, Sánchez & Baker, 2018, Teachers College Press).

Reviewers:
Marilyn Cochran-Smith, Boston College
Elizabeth Stringer Keefe, Lesley University
Wen-Chia Chang, Boston College
Molly Cummings Carney, Boston College

June 2018

Executive Summary

NCTQ recently released its 2018 review of U.S. teacher preparation programs. Employing open-records requests and online searches, the report ranks 567 graduate teacher preparation programs, 129 alternative route programs, and 18 residencies on practice, knowledge and admissions. The report seeks to determine if the teacher preparation programs are aligned with NCTQ’s standards. Such alignment, the report insists, will produce teachers “not only ready to achieve individual successes, but also to start a broader movement toward increased student learning and proficiency.” However, the report determines that most programs are not aligned with its standards. Accordingly, it finds “severe structural problems with both graduate and alternative route programs that should make anyone considering them cautious.” However, the report has multiple logical, conceptual, and methodological flaws. Its rationale includes widely critiqued assumptions about the nature of teaching, learning, and teacher credentials. Its methodology, which employs a highly questionable documents-only evaluation system, is a maze of inconsistencies, ambiguities, and contradictions. Further, the report ignores accumulating evidence that there is little relationship between graduates’ classroom performance and NCTQ’s ratings. Finally, the report fails to substantively account for broad shifts in the field of teacher education that are nuanced, hybridized, and dynamic. Regrettably, the report exacerbates the dysfunctional dichotomy between university programs and alternative routes and offers little guidance for consumers, policymakers, or practitioners.

Reviewers:
Marilyn Cochran-Smith, Boston College
Elizabeth Stringer Keefe, Lesley University
Wen-Chia Chang, Boston College
Molly Cummings Carney, Boston College

June 2018

I. Introduction

The National Council on Teacher Quality (NCTQ) is a private, non-profit advocacy organization, founded by the Fordham Foundation in 2001 to encourage alternative routes into teaching and to challenge the university “monopoly” on teacher education.1 Since 2013, NCTQ has ranked and rated teacher preparation programs according to a set of internal standards developed by the organization, and it has widely disseminated the results to state policymakers, top university leaders, and potential “consumers” of teacher preparation programs. NCTQ’s 2018 Teacher Prep Review,2 which is the focus of this review, is the latest in the organization’s controversial series of reports on the quality of teacher preparation programs.

II. Findings and Conclusions of the Report

Aimed directly at prospective “consumers” of teacher preparation programs, this report reviews 714 post-baccalaureate teacher preparation programs. These are labeled: “graduate” or “traditional” (n=567 programs); “alternative-route” or “ internship” (n=129 programs); and, “residency” (n=18 programs). The same scales were used in three areas to evaluate programs. These included: program requirements regarding practice prior to full classroom responsibility; knowledge, including candidates’ academic background, content knowledge, and pedagogical knowledge; and, program admissions criteria. The report presents key findings and makes recommendations in each of these areas. In sum, the report asserts that all programs should focus on and/or require NCTQ criteria for practice, knowledge, and admis-
sions.

The report also lists “top programs” (reported in terms of percentiles) for elementary teacher preparation (e.g., 99<sup>th</sup> percentile: INSPIRE Texas, YES Preparatory Public Schools Inc., Johns Hopkins University) and secondary teacher preparation (e.g., 99<sup>th</sup> percentile: CUNY-Hunter College, Richmond Teacher Residency, CUNY-Lehman College, Arizona State University InMAC Program and TFA Partnership, Teach for America DC, INSPIRE Texas).

Using input-based, internally developed standards to evaluate the quality of preparation programs based on syllabi and documents, the NCTQ report reaches the following conclusions:

### Practice

**• Student Teaching**

Only 6% of graduate and alternative-route programs give teacher candidates adequate practice with expert mentor teachers, observe candidates frequently enough, and provide adequate feedback. Residency programs do better with “about a third...paying attention” to these criteria.

**• Classroom Management**

“About half” of traditional graduate programs attend to “five research-based” classroom management strategies while “almost three quarters” of residencies and alternate programs do. Across programs, the most frequently required management strategies are establishing classroom rules/expectations and maximizing learning time; least frequently required of the five is teachers’ use of meaningful praise.

### Knowledge-Elementary Teacher Preparation

**• Elementary Content**

No graduate elementary programs adequately screen candidates for prior knowledge of elementary content, and only 15% ensure candidates gain this knowledge during the program. Thus 73% of graduate programs received a grade of D or F in this area. In contrast, a “sizable majority” of “a limited sample” of 28 alternative programs require passing scores on state content knowledge tests for program admission, thus meeting NCTQ’s knowledge standard. However, the report notes that the latter was the case because candidates became teachers of record while still in training.

**• Elementary Mathematics**
Only 1% of graduate elementary programs and “23 of 28” sample alternates adequately prepare elementary teachers in mathematics. The report concludes this “may stand as one of the most staggering weaknesses in teacher preparation, contributing to the chronically low standing of American schoolchildren in mathematics internationally.”

**Early Reading**

No alternate programs in the sample and only 23% of graduate programs provide adequate preparation.

**Knowledge-Secondary Teacher Preparation**

**Content in the Sciences**

Only 25% of graduate programs and 4% of alternative programs that offer certification in multiple science subjects ensure candidates have adequate content knowledge, primarily because they require “no more than a major” in one of these areas.

**Content in Social Studies**

44% of graduate programs and 25% of alternative programs offer certification in multiple social studies subjects ensure candidates have adequate content knowledge because they require “no more than a major” in one of these areas.

**Secondary Methods Coursework and Practice**

77% of graduate programs and 43% (of 80 alternative route programs) require teacher candidates to take a teaching methods course in their subject area.

**Admissions**

Only 14% of graduate programs and 23% of alternatives are rigorous enough in their selectivity criteria.

**III. Rationale for Findings & Conclusions**

The underlying rationale of this report is that prospective teachers who already have a bachelor’s degree in another area should have multiple teacher preparation options. This is a reasonable premise that is accepted more or less universally in the U.S. Beyond that, the report hinges on four assumptions. First, the report works from the assumption that teaching and
teacher preparation are primarily technical activities that can be defined, taught, and assessed by ticking off discrete bits of knowledge and skill, even in highly complex areas such as teaching reading or managing learning environments. Second, the report assumes that boosting human capital, especially teacher candidates’ academic credentials and content knowledge, is the key to enhancing teacher quality. Third, the report assumes that syllabi and document review is a valid measure of preparation program quality. Finally, the theory of action behind the report is that ranking and grading preparation programs will bring the force of the market to bear on teacher preparation, which is an effective catalyst for program reform and improvement.

IV. Report’s Use of Research Literature

There is very little research directly referenced in the report, and the large body of existing literature related to teacher policy and/or teacher preparation/licensure issues is mostly ignored. To support the report’s conclusions about teaching practice, including the assertion that all mentors should be pre-screened and all candidates frequently observed, the report relies on a single nearly 10-year old study, which is its only reference to a peer-reviewed article. It is problematic that the report does not account for the existing research on how varying features of clinical experience contribute to teacher candidates’ development or the emerging body of research on teacher residency programs and/or practice-based teacher education, especially given that the report lauds residency programs. With regard to classroom management, the report directs readers to two NCTQ documents, which detail “three authoritative summaries” on management, including an Institute of Education Science (IES) report. The wording in the NCTQ document implies that its “big five” of classroom management, which includes strategies such as establishing classroom rules and handling minor misbehavior without interrupting instruction, come straight from the IES report. But a close look reveals that “the big five” is NCTQ’s own construction, and the IES list of strategies is more elaborated, less technical, and more nuanced.

The report’s recommendations related to the knowledge elementary candidates need to teach math (e.g., “deep understanding” of the concepts underlying arithmetic) are purportedly based on what mathematicians recommend and what successful countries do. Few people would challenge the common sense idea that elementary teachers need to understand the math concepts they teach. But the report cites no specific research to support this claim. Instead, at the end of the entire math section, the report directs readers to a separate document regarding NCTQ’s elementary mathematics standard. That document includes what NCTQ describes as a “research inventory” comprised of 44 studies related to “the preparation of teacher candidates for elementary mathematics instruction.” According to NCTQ itself, however, not all the studies in the inventory are “directly relevant to the specific indicators of the standard, but rather they are related to the broader issues that the standard addresses” (p. 1). Thus it is unclear exactly whether and to what extent each study in the inventory is relevant to NCTQ’s standard and its recommendations regarding elementary teachers’ knowledge of math.
With regard to the teaching of reading, the report claims its criteria are based on “the best research available about what works in reading instruction.” This is the same problematic claim NCTQ reports have been making about reading instruction since 2006,9 a claim that has been critiqued and rejected by many experts in the field.10

Finally, regarding admissions standards, the report claims that “sixty years of research and evidence from nations whose students outperform American students demonstrate the benefits of teachers who have reasonably strong academic aptitude.” No evidence is provided for this claim. However, to get an “A” on admissions, preparation programs had to require a 3.0 individual GPA or a 3.3 cohort average along with GRE scores. Again no specific evidence is cited for this very specific requirement. Instead readers are again directed to a separate document,11 which cites 18 studies “investigating the selection of teacher candidates of high academic caliber.” As we have pointed out in two previous NEPC Reviews regarding NCTQ reports, the organization’s selection criteria misrepresent or ignore the nuanced findings of research on the purpose and impact of admissions test scores, which are not intended to predict teacher effectiveness and often co-vary with other factors.12 In addition, a recent analysis13 found no significant relationship between NCTQ’s selection criteria and teachers’ performance, as indicated by principals’ evaluations. This same analysis examined 13 studies cited by NCTQ as “strong evidence” for their admissions standard. However the reviewers found that most of NCTQ’s citations provided “little or no evidence to support the selectivity criteria, and many articles cited as evidence did not even address the topic.”14

V. Review of Report’s Methods

Although the report refers readers to a separate “methodology document,”15 the report’s methods are a maze of inconsistencies, ambiguities, and contradictions. In fact, to say that the report’s methods are confusing and unclear is an understatement.

Throughout the report and accompanying documents, there is inconsistency about “types” of programs reviewed, how many were reviewed, which standards were applied, how they were applied to which programs, and what the findings mean. All the programs reviewed in this report are post-baccalaureate. Unfortunately, clarity about the report’s methods ends there. On page 1, the report identifies three types of programs: traditional, alternative route, and residencies. While the report defines traditional programs as those “offered by colleges and universities,” it notes that alternative route programs can also be sponsored by colleges and universities, school districts, and other entities. This raises questions about how the report categorizes university-created alternative routes or, from a completely different angle, preparation programs offered at “new graduate schools of education,” which grant graduate degrees but are unaffiliated with universities.16 The report later refers to “pure” alternative programs, implying two types—residencies wherein candidates work with a mentor for a year and “pure” alternatives that “plunge” candidates into full-time responsibility with or without support. This bifurcation further muddies the waters and does not acknowledge that many residencies are supported by universities in partnership with schools, as evidenced by the fact that the vast majority of members of the National Center for Teacher Residencies are
university-initiated or university-connected residency programs.\textsuperscript{17}

The report’s analyses exacerbate the confusion about program types. When referring to some areas of student teaching and classroom management, the report lumps alternative programs and residencies into one category, while in other places, the three “types” are reported separately. Especially problematic is that in some discussions, such as those regarding content/method standards, most of the graphs provided portray only university programs, noting in the text that “a sample” of 28 alternative route programs was reviewed. No detail is provided about what the 28 programs were, or how and why only a sample was selected, or whether they were “pure” or “residency” programs. It is difficult to sort teacher preparation programs using dichotomous labels (traditional and alternative) as this report seems to want to do, despite the fact that it begins with three types. For years, it has been concluded that the labels “traditional” and “alternative” are largely meaningless, given that there is more variation among than between these groups.\textsuperscript{18}

The report’s use of input-based, internally developed standards to evaluate the quality of programs based on syllabi and documents has been controversial and widely critiqued since the method was first proposed in the early 2000s.\textsuperscript{19} We do not rehash the critiques here. However we do note that it has been strenuously (and we would say, persuasively) argued that it is not possible to produce valid, accurate, and usable assessments of the quality of teacher preparation programs (or courses) based on syllabi and documents alone without also gathering other data through participant and faculty surveys, site visits, and other means that get at what programs actually do and how participants actually experience them.\textsuperscript{20} The salience of these difficulties is illustrated in the report’s surprising use of criteria that include the quality of lectures and the nature and quality of the feedback supervisors provide to student teachers, both of which would seem to require information unavailable in written documents.

The report asserts that all programs were “ranked on the same scale.” This is especially confusing given that the methodology document\textsuperscript{21} to which readers are referred has two sets of standards, one for “traditional” programs with 19 standards and one for “alternative” programs with five standards. Supplementary NCTQ documents\textsuperscript{22} indicate that some standards for traditional programs were also applied to alternative programs, but it is not clear which standards exactly were used to evaluate the report’s “limited” sample of alternative programs.

For overall rankings, not all standards were used. According to the report, elementary, secondary, and special education program rankings were based on “key” and “booster” standards,\textsuperscript{23} with scores for certain standards weighted more than others. For elementary programs, admissions criteria were weighted heaviest, followed by student teaching; content areas were weighted less. For secondary programs, admissions criteria were weighted heaviest, followed by content and student teaching. Program rankings were “boosted” by scores on the classroom management standard or, for secondary programs, the methods standard.
There is no information provided in the report about the rationale for the weightings, how “heavier” or “lesser” weightings were calculated, or how scores on “booster” standards were figured into overall rankings.

Individual “trained” raters used “low inference” protocols to answer “yes or no” as to whether syllabi and documents included references to the specific information or techniques in question. Based on these, a standard grade between “A” (“meets the standard”) and “F” (“does not meet the standard”) was “automatically generated.” Some standards (e.g. student teaching) were ranked on a five-point scale (A, B, C, D, or F), while some (e.g., candidate selection) were ranked on a three-point scale (A, C, or F). It is not clear why. About 20% of programs had a second reviewer, but rater reliability estimates are not reported.

Top programs are reported in terms of percentiles, which presumably means that if a program is at the 99th percentile, its score is higher than the scores of 99% of programs in the entire pool. This makes sense statistically only if every elementary and secondary program is evaluated by the same set of standards. The fact that it is unclear which standards were used for all programs and which alternative programs were evaluated using the “complete set of key standards” makes cross-program percentile rankings potentially problematic.

VI. Review of Validity of Findings and Conclusions

Some of the report’s recommendations regarding the improvement of post-baccalaureate teacher preparation programs are reasonable. For example, there is research that supports the report’s recommendations that all teacher candidates should have rich classroom experience with quality feedback from knowledgeable mentors prior to taking on full teaching responsibility. However, the overall validity of the report’s findings and conclusions is fundamentally contingent on the validity of its rationale, use of research, and methods. The report’s rationale is based on the questionable assumption that teaching, learning, and learning to teach are primarily technical activities even in highly complex areas such as teaching reading or managing learning environments. In addition, the report assumes that syllabi and document review is a valid measure of program quality despite widespread challenges to this approach and accumulating evidence that there is little relationship between the performance of teacher preparation program graduates and their programs’ NCTQ ratings. Further, as we note above, the report ignores most of the related existing research, and many of the report’s references to research are misleading. Finally, it is highly unlikely that any other group of researchers could replicate the report’s methods and reach similar conclusions, given multiple inconsistencies, ambiguities, and incomplete or contradictory information about methodology. The report’s stated goals to the contrary, what has been produced in this report is not a reasoned and fair analysis of preparation program quality but a confusing and idiosyncratic analysis of a highly selective set of artifacts. Despite far-reaching critiques, the report treats its rationale as self-evident, and NCTQ continues to produce teacher preparation reports that are misleading, confusing, and flawed logically, conceptually, and methodologically.
VII. Usefulness of the Report for Guidance of Policy and Practice

Although this *Teacher Prep Review* tackles less information than in previous NCTQ reviews, it lacks cohesion and is user-unfriendly in the extreme. There are inconsistencies and questionable conclusions throughout the report, making it disjointed and hard to follow. From its poorly executed narrative-style introduction to its confusing graphics, distracting typos, and incomplete description of research methods, the report offers little to consumers seeking information about program quality or to program leaders interested in improvement. Further, the report fails to account for ongoing shifts in teacher education as an institutional field, which is considerably more nuanced, hybridized, and dynamic than the report implies. In the end, even though the report warns consumers to treat all post-baccalaureate programs with caution, it functions primarily to exacerbate the dysfunctional dichotomy between university programs and alternative route programs and offers little guidance for policymakers, practitioners, or consumers.
Notes and Resources


3 See, for example:


5 See, for example:

http://nepc.colorado.edu/thinktank/review-teacher-prep-2018


See, for example:


18 See for example:


20 See for example:


25 See, for example:


26 See, for example:


in the United States. *The Educational Forum, 77*(1), 6-27;


See, for example:


Dudley-Marling. (2013, September 1). Re: *Teacher prep gets failing grade: Study faults education schools, but the schools say the study is badly flawed* [Article comment]. Retrieved January 25, 2015, from: http://commonwealthmagazine.org/education/003-teacher-prep-gets-failing-grade/;


http://nepc.colorado.edu/thinktank/review-teacher-prep-2018 16 of 16