Research inventories on features of teacher preparation relevant to NCTQ standards: Rationale and methods

Rationale

NCTQ's research inventories serve a dual purpose. First, by continuing to be apprised of research on teacher preparation, we ensure that our standards are based on the best available evidence of what makes new teachers effective in the classroom. Second, the inventories provide the big picture on the kind and quantity of research on teacher training. Unfortunately, the big picture currently reveals a dearth of actionable, gold-standard research in the field, despite the fact that a database search on "teacher preparation" or "teacher education" yields approximately one million articles.

Over 3,000 relevant peer-reviewed journal articles were evaluated for consideration in our inventories, with approximately 365 ultimately considered relevant to teacher preparation.[1] Of these 365 articles, we determined that fewer than 30 (eight percent) are both well designed and bear *directly* on the issue of how teacher preparation affects teacher effectiveness.

Similar research winnowing processes have been conducted recently, all with similar results. Arthur Levine wrote in *Educating School Teachers:* "[T]here has been a dearth of systematic research documenting the impact of teacher education programs on the students their alumni teach."[2] Levine cites the following:

- A Michigan State University review of teacher education research that examined the scholarly works published in three broad areas of teacher education: 1) subject matter preparation, pedagogical education and clinical training needed by prospective teachers; 2) the policies and strategies used successfully to improve and sustain the quality of preservice teacher education; and 3) the characteristics of high-quality alternative certification programs. The authors limited their review to research on teacher education in the United States published over the preceding 20 years in peer-reviewed journals. This produced a total of 313 articles, only 57 of which meet generally accepted standards of research.[3]

- A 2005 meta-study by a panel of the American Educational Research Association (AERA) that culminated in an 804-page report filled with expressions such as "so few studies" (p. 427), "sobering to look at amount of empirical research done" (p. 282), "extremely thin" (p. 287), "uneven" (p. 600), "limited" (p. 26), "so little existing research" (p. 619), "we know next to nothing" (p. 610), "relatively few empirical studies" (p. 651), "very few studies were longitudinal" (pp. 489-90), "vagueness of criteria for evaluation" (p. 674) and "almost nonexistent"(p. 27).[4]

---

[1] No inventories were developed for Standard 17 (Outcomes) and Standard 18 (Impact on Student Learning). In the case of the former, the standard relates to institutional features of preparation, not to preparation itself. In the case of the latter, the standard itself bears directly on teacher effectiveness.

[2] Levine, A. (September 2006). *Educating School Teachers* (p. 19). Washington, DC: The Education Schools Project.
What do the majority of studies in teacher preparation examine if they are not focused on teacher effectiveness? The majority of studies instead investigate the perceived self-efficacy of preservice teachers. Since there is no evidence that perceived self-efficacy bears any relationship to effectiveness, this research can hardly be a guide to how to best prepare teachers.

[3] Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps and recommendations.* Seattle, WA: University of Washington, Center for the Study of Teaching and Policy. Retrieved from http://depts.washington.edu/ctpmail/PDFs/TeacherPrep-WFFM-02-2001.pdf#search=%22the%20teacher%20preparation%20research%3A%20current%20knowledge%2C%20gaps%22

[4] Cochran-Smith, M., & Zeichner, K. M. (Eds.). (2005). *Studying teacher education: The report of the AERA panel on research and teacher education.* Hillsdale, NJ: Lawrence Erlbaum Associates.

- A literature review conducted in 2000 by SRI International for the Clinton administration characterized the scholarship in teacher education as "not particularly robust" and went on to say: "The evaluative frame of mind has not yet penetrated teacher education."[5]

The absence of a research-based foundation for teacher preparation has frustrated even the most concerted efforts to distill best practices. In 2004, Congress directed the National Research Council (NRC) to conduct a study of teacher preparation programs in the United States. In its report, the NRC concluded that research on teacher preparation "includes a relatively small body of empirical studies that provide some evidence about the effects of particular kinds of instruction; it also includes an even smaller amount of evidence about the effects of particular approaches to teacher preparation."[6]

Methods

Selection process

Our search encompassed peer-reviewed journal articles that describe research on the preparation of preservice teachers for K-12 education for which the full text in English was available through the ERIC, EBSCOHOST, JSTOR, EconLit, PsychArticles or PsychInfo databases. We specifically looked at research areas pertaining to NCTQ standards, conducting multiple searches using a variety of search terms in different combinations. Not all studies in a Research Inventory are immediately relevant to an NCTQ standard. Rather, they encompass the research in the field that has some general bearing on broad topics of teacher preparation (e.g., elementary content or student teaching), but not all studies directly relate to the indicators within NCTQ's standards (e.g., the specific topics elementary teachers must know or the number of times a student teacher should be observed). However, the studies meeting the criteria for strong design helped inform the standards and indicators for the *Teacher Prep Review*.

To develop these inventories, all relevant research published in at least the last five years was considered. In most inventories, research published in the last 10 years was considered. (The time period relevant to each inventory is noted in the inventory itself.) Research that only tangentially addresses specific methods of preservice preparation was not included. Some articles that focus on in-service or alternate route programs, however, are included if the relevant features could be replicated in a preservice preparation setting.

Classification

Research was classified in two stages: first considering design "strength" relative to several variables common to research designs, and second, considering whether student effects (as measured by external, standardized assessments) were considered.[7]

"Studies with stronger design" use some sort of control or comparison group in an experiment, natural or otherwise, or use a multiple regression for evaluation. These studies have a sample size of 100 or more unless the subjects involved are not individuals (e.g., teacher preparation programs) in which case the minimum sample size was determined based on the context of the study and the nature of the subjects. In the case of experiments, the number of subjects in each of the treatment and control groups had to total 100 or more to classify the relevant study as having "strong design." In cases in which dyadic groups were analyzed, 50 participants constituted the minimum sample size for categorization as having "strong design."

---

[5] Humphrey, D. C., Adelman, N., Esch, C., Riehl, L., Shields, P. M., & Tiffany, J. (2000). *Preparing and supporting new teachers: A literature review* (p. 17). Washington, DC: SRI International, U.S. Department of Education.
The teacher educator community has not been oblivious to its lack of what SRI International so felicitously described as an "evaluative frame of mind." Fifteen years ago a professor at a New York school of education coauthored an article with the title: "Measurable Change in Student Performance: Forgotten Standard in Teacher Preparation." Greenwood, C. R., & Maheady, L. (1997). Measurable change in student performance: Forgotten standard in teacher preparation? *Teacher Education and Special Education, 20*(3), 265-275.

[6] National Research Council. (2010). *Preparing teachers: Building evidence for sound policy* (p. 175). Washington, DC: National Academy Press.

[7] NCTQ will revisit these research classifications periodically to reflect new study design criteria established by the What Works Clearinghouse and other similarly rigorous research and evaluation entities.

“Studies with weaker design” have no comparison or control, can have subjects whose relationship to the researcher creates bias, are case studies with potential selection bias, or rely on survey or otherwise qualitative data. These studies often have a sample size of fewer than 100.

Some studies with control groups were categorized as having weak design due to the issues just mentioned or if the control group was inappropriately selected or the study did not provide enough details about the control group to rule out significant differences between the treatment and control groups.

In the case of studies that had both strong and weak characteristics, categorization was determined by whether the research would be useful for teacher educators, teacher education program administrators and/or policymakers. If it seemed potentially useful, it was categorized as “strong design.”

NCTQ welcomes additions or suggested corrections to the inventories by experts in the field.