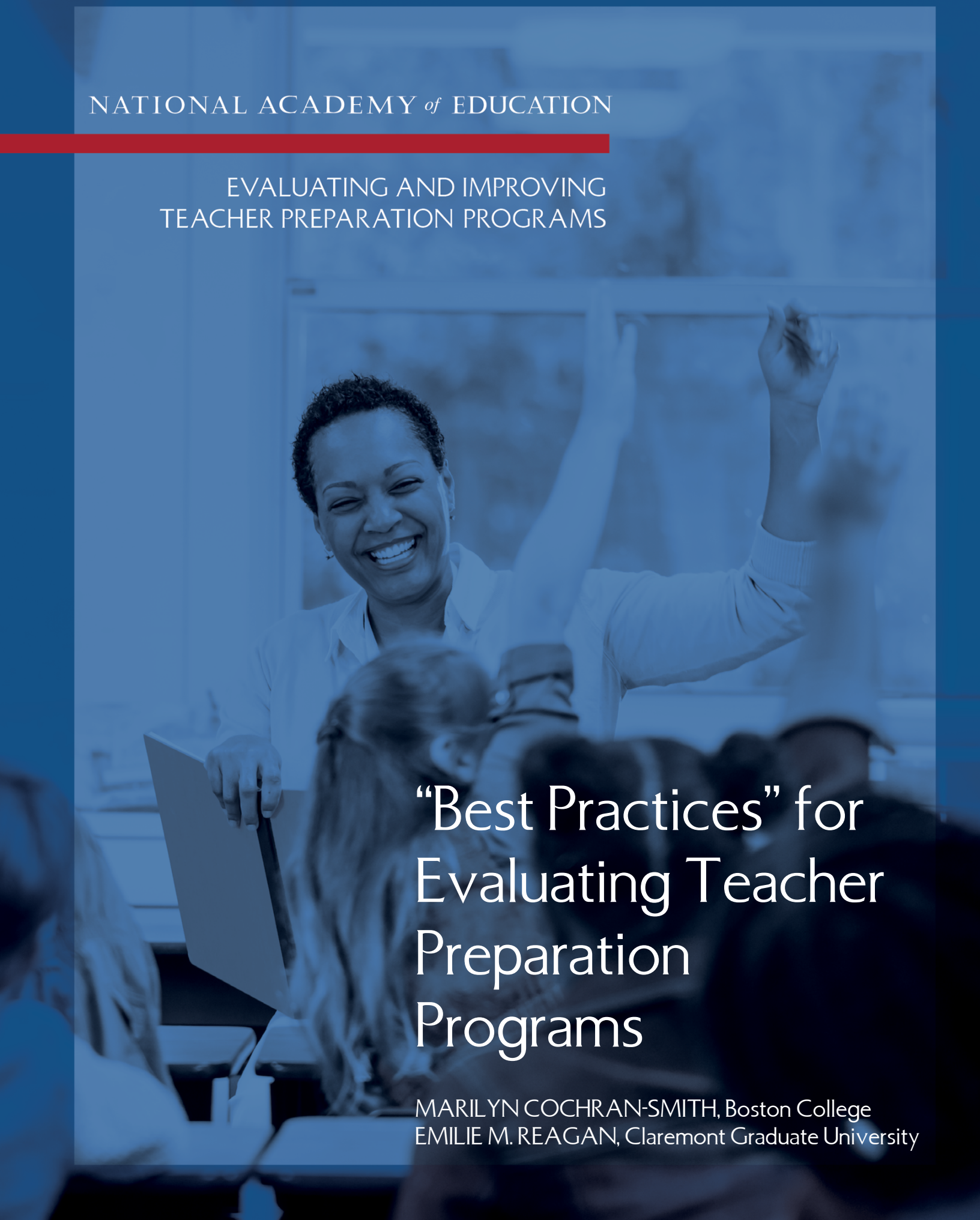
NATIONAL ACADEMY of EDUCATION

EVALUATING AND IMPROVING
TEACHER PREPARATION PROGRAMS



"Best Practices" for Evaluating Teacher Preparation Programs

MARILYN COCHRAN-SMITH, Boston College
EMILIE M. REAGAN, Claremont Graduate University

NATIONAL ACADEMY OF EDUCATION 500 Fifth Street, NW Washington, DC 20001

NOTICE: The project and research are supported by funding from the Bill & Melinda Gates Foundation. This paper was prepared for the National Academy of Education (NAEd) to inform and support the work of the steering committee for *Evaluating and Improving Teacher Preparation Programs*, including the consensus report. The opinions expressed are those of the authors and not necessarily those of the NAEd or the Bill & Melinda Gates Foundation.

Digital Object Identifier: 10.31094/2021/3/2

Suggested citation: Cochran-Smith, M., & Reagan, E. M. (2021). *"Best practices" for evaluating teacher preparation programs*. National Academy of Education Committee on Evaluating and Improving Teacher Preparation Programs. National Academy of Education.

National Academy of Education

Evaluating and Improving Teacher Preparation Programs

"Best Practices" for Evaluating Teacher Preparation Programs

Marilyn Cochran-Smith, *Boston College*
Emilie M. Reagan, *Claremont Graduate University*

September 2021

CONTENTS

INTRODUCTION

In 2010, as the result of a congressionally mandated study, the National Research Council (NRC) published the report *Preparing Teachers: Building Evidence for Sound Policy* (NRC, 2010). Reflecting the unprecedented attention to teacher quality that had emerged internationally in response to the exigencies of the "global knowledge economy" (McKinsey & Company, 2007; Organisation for Economic Co-operation and Development, 2005; World Bank, 2010), the report was intended to respond to policymakers' demands to know the extent to which the characteristics, practices, and policies that typified teacher preparation in the United States were or were not consistent with scientific evidence. The NRC report reached three key conclusions: that there was enormous variation both between and within differing pathways into teaching, rather than one clearly superior route; causal evidence linking characteristics of teacher candidates and/or preparation programs with student achievement or other outcomes was complex and very difficult to develop; and that there was a need for a comprehensive data collection system in the United States that would support quality control and accountability in teacher preparation.

Since 2010, there have been multiple reports and other documents that propose recommendations about how to evaluate, assess, or hold teacher preparation accountable and/or about how to use evaluation information to improve teacher preparation. As the authors of this paper, we were charged with preparing an analysis of recent work regarding "best practices for evaluating teacher preparation programs" by synthesizing and critiquing major reports explicitly focused on teacher preparation evaluation. To fulfill this charge, we reviewed 19 major reports about teacher preparation evaluation, assessment, or accountability published between 2010 and 2020. Our analysis revealed that the primary goal of the majority of existing reports was identifying the strengths and weaknesses of evaluation metrics based on rigorous criteria for accuracy and utility. Our analysis also revealed that the majority of reports did not position equity as a central goal of evaluation and actually said very little about equity explicitly, although some assumed that equity was a by-product of rigorous evaluation systems.

Building on our analysis of the 19 reports, this paper calls for a new equity-centered approach to teacher preparation evaluation that acknowledges the serious inequities in educational opportunity and attainment across groups in the United States as well as the important role that teacher preparation evaluation can play as part of larger efforts to overcome disparities in opportunity and attainment. We argue that strong equity, which we elaborate on below, should be established as an explicit goal and a desired outcome of teacher preparation evaluation, and that it should be central to the design, interpretation, uses, and consequences of evaluation.

TEACHER PREPARATION EVALUATION: A COMPLEX LANDSCAPE

We begin with a brief overview of the complex landscape related to teacher preparation evaluation. Lack of consensus about the value of teacher preparation, coupled with market-based responses to the perceived pressures of the global economy (Ambrosio, 2013; Scott, 2016), have combined with other forces over the past three decades to produce a crowded, rapidly changing, and fragmented teacher education field (Lincove et al., 2015) characterized by competing reform agendas (Cochran-Smith & Fries, 2001;

Zeichner, 2018). Within this larger context, by the early 2000s, there was widespread attention to teacher preparation evaluation and accountability from both within and outside the field. In fact, accountability was regarded by many policy and other actors as a key mechanism for "fixing" teacher preparation, which was characterized as a "broken" profession and a "broken" system (Duncan, 2009; U.S. Department of Education, 2002, 2003). Understanding the landscape of teacher preparation evaluation involves sorting out the intersecting efforts and sometimes conflicting roles of state and federal agencies, philanthropies, independent advocacy organizations, and professional accreditors and organizations.

Federal and State Roles in Teacher Preparation Evaluation

In the early part of the 2010s following the passage of Race to the Top legislation in 2009, the federal government issued a bold new blueprint for the reform of teacher preparation—*Our Future, Our Teachers* (U.S. Department of Education, 2011). This report was consistent with the gradual shift in education accountability in the United States from local to state and federal levels that had been occurring since the mid-1980s (Bales, 2006); it was also consistent with the Obama administration's education reform agenda, which, building on the efforts of the previous administration, relied on market competition to elevate good programs and drive bad programs out (Au, 2016; Lipman, 2011; Scott, 2016; Taubman, 2009). The blueprint aimed to tie federal resources to the achievement of the students taught by graduates of identifiable teacher preparation programs and pathways, thus connecting the federal, state, and institutional policy levels (Cochran-Smith et al., 2013). Although the 2011 blueprint later died in committee, many of its policies were resurrected in a starker form in the Title II Higher Education Act (HEA) regulations proposed in 2014 (U.S. Department of Education, 2014), which stipulated that existing inputs-oriented annual reporting regulations be replaced by outcomes-oriented measures of student achievement, graduates' job placement and retention data, and graduates' and principals' program satisfaction data (Cochran-Smith et al., 2018). Although the 2014 proposal had broad bipartisan political support, the proposal prompted unprecedented public and professional opposition extending over almost 2 years. Nevertheless, the new regulations were approved in the last months of the Obama administration (U.S. Department of Education, 2016b), then almost immediately rescinded by the Trump administration (Iasevoli, 2017).

At the state level, over the course of the 2010s, policymakers and state education agencies continued efforts to improve state approval requirements for teacher preparation, with a similar shift in many states toward outcomes-based accountability. States that were awarded Race to the Top grants were required to develop data systems linking preparation programs to K-12 student achievement using growth modeling and value-added assessments, a trend followed by some other states (Von Hippel & Bellows, 2018). Additionally, in 2012, the Council of Chief State School Officers (CCSSO) created a multi-state, multi-year reform effort, the Network for Transforming Educator Preparation (NTEP), to leverage state authority over preparation program approval and licensure, with data systems being a key lever (CCSSO, 2018).

Despite these developments, as the 2010s went on, there were challenges to state-level data systems, the withdrawal of broader federal policy levers, and growing evi-

dence questioning the validity, reliability, stability, and utility of inferences based on value-added measures and growth modeling for the purpose of evaluating individual teachers and/or teacher preparation programs (e.g., American Education Research Association, 2015; Haertel, 2013; Noell et al., 2019). At the same time, drawing on evidence regarding the potential of performance assessments for evaluating teaching practice (Darling-Hammond et al., 2012), many state education agencies adopted nationally available or state-initiated performance assessments (e.g., edTPA, Educational Teaching Service National Observational Teaching Examination, Massachusetts Candidate Assessment of Performance) as a requirement for teacher certification and/or program approval (AACTE, 2020).

Philanthropic and Advocacy Group Involvement in Teacher Preparation Evaluation

During this same time, there was continued philanthropic interest in teacher preparation evaluation. Hess (2005) called these efforts "muscular philanthropy," or, large gifts funded by a small group of donors (e.g., S. D. Bechtel, Jr. Foundation, Bill & Melinda Gates Foundation, Charles and Lynn Schusterman Family Philanthropies, The Eli and Edythe Broad Foundation, NewSchools Venture Fund, Walton Family Foundation) tied to expectations about disruption, innovation, and accountability (Colvin, 2005; Zeichner & Peña-Sandoval, 2015). For example, multiple philanthropies funded private advocacy organizations such as the controversial National Council on Teacher Quality (NCTQ), which critics have excoriated because of their highly-politicized report cards for teacher preparation programs based on the organization's own criteria, which were never vetted by the profession (Cochran-Smith et al., 2013, 2018; Fuller, 2015). Another example of the increased role of philanthropies in teacher preparation evaluation is the Bill & Melinda Gates Foundation's Teacher Preparation Transformation Centers, which funded TPI-US, an independent inspectorate, to review teacher preparation programs with the purpose of highlighting practices worthy of expansion (TPI-US, 2020). As of 2021, TPI-US has completed more than 150 inspections of teacher preparation programs in 21 states, and it has state-level partnerships in Florida, Louisiana, and Massachusetts. Here it is worth noting that the National Academy of Education (NAEd) project on *Evaluating and Improving Teacher Preparation Programs*, of which this paper is part, was funded by the Bill & Melinda Gates Foundation, and many of the reports analyzed herein were funded by the philanthropies listed above. Zeichner & Peña-Sandoval (2015) have suggested that efforts like these represent an "outsized role" of private interests in teacher preparation policy.

Professional Involvement in Teacher Preparation Evaluation

With the 2010s came major shifts in national teacher preparation programmatic accreditation, reflecting a lack of consensus about evaluation within the profession. In 2013, the two existing accrediting bodies, the National Council for Accreditation of Teacher Education (NCATE) and the Teacher Education Accreditation Council (TEAC), merged to form the Council for the Accreditation of Educator Preparation (CAEP), with the goal of presenting a unified voice and elevating the status of the profession. Building

on federal HEA Title II reporting regulations, CAEP's standards required preparation programs to demonstrate candidates' and graduates' impact on K-12 student learning, recommending value-added approaches and growth modeling of student achievement. There was enormous controversy surrounding these outcome standards and about candidate selectivity standards, which threatened CAEP's credibility within and outside the profession (Cochran-Smith et al., 2018). In 2017, a new national accrediting body, the Association for Advancing Quality in Educator Preparation (AAQEP), was founded, partly in response to critiques of CAEP. Describing itself as a quality assurance organization, AAQEP (2020) tied accreditation to innovation, quality, and responsiveness to program context, explicitly stating that it was grounded in trust of the profession and with standards developed collaboratively with stakeholders. As of February 2021, AAQEP had accredited more than 25 programs, and currently has more than 120 member institutions (AAQEP, 2020).

While there was disagreement within the profession about national programmatic accreditation, there was general convergence about the importance of clinical practice (AACTE, 2018; NCATE Blue Ribbon Commission, 2010). Many preparation programs implemented some version of practice-based teacher education, including clinically rich teacher preparation, teacher residency programs, and/or emphasis on "core" practices. Along these lines, there was increased attention to the development of measures linking clinical experience and teaching practice (e.g., performance assessments), the quality of program and K-12 school partnerships, and the "effectiveness" of cooperating teachers and field-based teacher educators (e.g., Goldhaber, 2019; Goldhaber et al., 2019; Ronfeldt, 2021; Ronfeldt et al., 2018).

The Role of Equity Agendas in Teacher Preparation Evaluation

During the 2010s, there were many efforts by preparation programs and by some professional collaborations to make equity and social justice the centerpiece of teacher preparation programs. There were also excoriating critiques of racial injustice within the larger field of teacher education itself and its general failure to acknowledge and respond to its own history of White supremacy (e.g., Anderson, 2019; Andrews et al., 2019a; Brown, 2013; Daniels & Varghese, 2019; Horn & Kane, 2019; Milner et al., 2013; Philip et al., 2019; Sleeter, 2017; Souto-Manning, 2019; Stillman & Beltramo, 2019). These criticisms built on a long history of critique by scholars who had advocated over many years for teacher education to address head-on issues of culture, race, social justice, equity, and the values of minoritized groups in the curriculum, fieldwork, policy, and practice (e.g., Cochran-Smith, 1995a, 1995b, 2010; Gollnick, 1992; Grant, 2008; King, 2008; Ladson-Billings, 1999 Nieto, 2010; Sleeter, 2001, 2009; Villegas & Lucas, 2004; Zeichner, 2003, 2009; Zeichner & Hoeft, 1996). Despite historical and contemporary critiques, however, as we elaborate in later sections of this paper, during teacher education's "era of accountability" from roughly 1998-2018 (Cochran-Smith et al., 2018), there was little explicit attention to equity as a goal of evaluation.

METHODS AND ANALYTIC FRAMEWORK

In this paper, we analyze and critique major reports, published between 2010 and 2020, whose explicit topic is the nature, characteristics, and/or strengths and limitations of teacher education evaluation, assessment, and accountability systems, tools, or initiatives in the United States. To constitute this body of literature, we used the search terms "teacher education" (or "teacher preparation" or "teacher quality") *and* "evaluation" (or "assessment" or "accountability") to locate reports, books, policy briefs, and other documents published by relevant professional organizations or by peer-reviewed academic presses. We also solicited suggestions for additional reports that met our criteria from senior scholars in the field and from the NAEd commissioning committee.

In gathering this body of work, we did *not* include reports focused on *teacher* rather than *teacher preparation* evaluation, nor did we include reports about educational evaluation in general rather than about teacher preparation explicitly. Furthermore, because of overlap with other papers in the NAEd commissioned paper series, we excluded both reports about particular evaluation tools (e.g., teacher performance assessments) and analyses of particular states' evaluation systems. Based on this search process, we identified 19 reports and other documents that met our criteria, with several authors and organizations producing more than one report. Table 1 includes an alphabetical

TABLE 1 Teacher Education Evaluation/Assessment/Accountability, Major Reports, 2010-2020

Report	Year	Publisher/Sponsoring Organization
Allen, M., Coble, C., & Crowe, E. (2014). *Building an evidence-based system for teacher preparation.* Teacher Preparation Analytics.	2014	Teacher Preparation Analytics
Almy, S., Tooley, M., & Hall, D. (2013). *Preparing and advancing teachers and school leaders: A new approach for federal policy.* The Education Trust. https://edtrust.org/wp-content/uploads/2013/10/Preparing_and_Advancing_0.pdf.	2013	The Education Trust
Carinci, J. E., Meyer, S. J., & Jackson, C. (Eds.). (2020). *Linking teacher preparation program design and implementation to outcomes for teachers and students.* Information Age Publishing.	2020	Information Age Publishing
Cochran-Smith, M., Carney. M. C., Keefe, E. S., Burton, S., Chang, W. C., Fernandez, M. B., Miller, A. F., Sanchez, J. G., & Baker, M. (2018). *Reclaiming accountability in teacher education.* Teachers College Press.	2018	Teachers College Press
Cochran-Smith, M., Stern, R., Sánchez, J. G., Miller, A., Keefe, E. S., Fernández, M. B., Chang, W. C., Carney, M. C., Burton, S., & Baker, M. (2016). *Holding teacher preparation accountable: A review of claims and evidence.* National Education Policy Center. https://nepc.info/sites/default/files/pb_cochran-smith_teacher_prep_0.pdf.	2016	National Education Policy Center
Coggshall, J. D., Bivona, L., & Reschly, D. J. (2012). *Evaluating the effectiveness of teacher preparation programs for support and accountability.* National Comprehensive Center for Teacher Quality. https://files.eric.ed.gov/fulltext/ED543773.pdf.	2012	National Comprehensive Center for Teacher Quality
Council of Chief State School Officers. (2012). *Our responsibility, our promise: Transforming educator preparation and entry into the profession*. Task Force on Educator Preparation and Entry into the Profession. https://www.ccsso.org/sites/default/files/2017-10/Our%20Responsibility%20Our%20Promise_2012.pdf.	2012	Council of Chief State School Officers

TABLE 1 Continued

Report	Year	Publisher/Sponsoring Organization
Council of Chief State School Officers. (2017). *Transforming educator preparation: Lessons learned from leading states*. https://ccsso.org/sites/default/files/2017-11/CCSSO%20Educator%20Preparation%20Playbook.pdf.	2017	Council of Chief State School Officers
Council of Chief State School Officers. (2018). *Measuring what matters: Recommendations from states in the network for transforming educator preparation*. http://www.ccsso.org/sites/default/files/2018-02/Measuring%20What%20Matters.pdf.	2018	Council of Chief State School Officers
Crowe, E. (2010). *Measuring what matters: A stronger accountability model for teacher education*. Center for American Progress. https://files.eric.ed.gov/fulltext/ED518518.pdf.	2010	Center for American Progress
Crowe, E. (2011a). *Race to the Top and teacher preparation analyzing state strategies for ensuring real accountability and fostering program innovation.* Center for American Progress. https://cdn.americanprogress.org/wp-content/uploads/issues/2011/03/pdf/teacher_preparation.pdf?_ga=2.1926006.694985793.1608421266-1212376878.1604456340.	2011	Center for American Progress
Crowe, E. (2011b). *Getting better at teacher preparation and state accountability strategies, innovations, and challenges under the federal Race to the Top program.* Center for American Progress. https://cdn.americanprogress.org/wp-content/uploads/issues/2012/01/pdf/teacher_preparation_execsumm.pdf.	2011	Center for American Progress
Deans for Impact. (2016). *From chaos to coherence: A policy agenda for accessing and using outcomes data in educator preparation*. http://www.deansforimpact.org/wp-content/uploads/2016/11/From_Chaos_to_Coherence.pdf.	2016	Deans for Impact
Education Deans for Justice and Equity. (2019a). *Seven trends in U.S. teacher education and the need to address systemic injustices*. National Education Policy Center. https://nepc.colorado.edu/sites/default/files/publications/FYI%20Ed%20Deans_1.pdf.	2019	Education Deans for Justice and Equity
Education Deans for Justice and Equity. (2019b). *A framework for assessment and transformation* ("The EDJE Framework" v.6.20.2019). https://drive.google.com/file/d/1AVVkmzPxEMAt_g-BPYyNVHgseylmREnN/view.	2019	Education Deans for Justice and Equity
Feuer, M. J., Floden, R. E., Chudowsky, N., & Ahn, J. (2013). *Evaluation of teacher preparation programs: Purposes, methods, and policy options.* National Academy of Education. https://naeducation.org/wp-content/uploads/2016/11/028489-Evaluation-of-Teacher-prep.pdf.	2013	National Academy of Education
Kumashiro, K. K. (2015). *Review of proposed 2015 federal teacher preparation regulations.* National Education Policy Center. https://nepc.colorado.edu/sites/default/files/ttr10-tchrprepregs_0.pdf.	2015	National Education Policy Center
National Research Council. (2010). *Preparing teachers: Building evidence for sound policy*. The National Academies Press.	2010	National Research Council
Worrell, F., Brabeck, M., Dwyer, C., Geisinger, K., Marx, R., Noell, G., & Pianta, R. (2014). *Assessing and evaluating teacher preparation programs*. American Psychological Association. https://www.apa.org/ed/schools/teaching-learning/teacher-preparation-programs.pdf.	2014	American Psychological Association

list of these reports with bibliographic information. Below, we discuss the analytic framework that guided our review of the reports and our positionality as the authors of this paper.

Analytic Framework: Theories of Evaluation

As we note above, over the past two decades, a growing number of professional organizations, philanthropies, consultants, advocacy organizations, and academic groups have made recommendations about how evaluation should be done and how accountability systems should operate in teacher preparation. These multiple organizations do not necessarily work from the same assumptions about the purpose of teacher preparation or the nature and overall purposes of teacher preparation evaluation. They also disagree about the best measures and metrics to use, who should be included as relevant stakeholders, and what the appropriate roles and relationships of stakeholders and evaluators should be.

Given these differences and in order to synthesize and critique the 19 reports we identified, we first organized them according to their underlying theories or models of evaluation. To do so, we drew on well-known frameworks for describing the history and landscape of the cross-disciplinary field of program evaluation. In seminal work in this area, Alkin and Christie (2004; Christie & Alkin, 2008, 2013) identified three major approaches or models of evaluation, which they labeled *methods*, *use*, and *valuing*. A few years later and building on Alkin and Christie's work, Mertens and Wilson (2012, 2019) suggested four paradigms of evaluation that roughly mapped onto, but also extended, Alkin and Christie's models, which they labeled *post-positivist*, *pragmatic*, *constructivist*, and *transformative*. We drew on these frameworks to identify three approaches to teacher preparation evaluation, which we use to organize, synthesize, and critique the 19 reports. It is important to note here that these categories were developed to organize and facilitate analysis across dozens of evaluation theorists' espoused prescriptions of how evaluation should be practiced. Particular evaluations and evaluators may, and often do, incorporate assumptions from multiple models, blending and adapting approaches in practice rather than tightly adhering to a particular approach (Alkin & Christie, 2019). However, when considering trends and characteristics across evaluations, as we do in this paper, this framework is a useful heuristic for identifying and contrasting differing assumptions and theories of action.

Post-Positivist, Methods-Focused Approaches to Evaluation

According to Alkin's (2004) seminal volume, the major purpose of evaluation is to assess the degree to which programs are accountable for their actions and use of resources coupled with the public desire for systematic and justifiable *methods* of accountability in keeping with the conventions of social inquiry. As Alkin and Christie point out (Alkin, 2004; Alkin & Christie, 2004), Cook and Campbell (1979) were central in defining this perspective on evaluation *as* research wherein high-quality evaluation depends on the application of rigorous research methods with the goal of producing generalizable findings. This perspective is generally consistent with what Mertens and Wilson (2012, 2019) refer to as the "post-positivist paradigm" in evaluation. This

paradigm recognizes that although knowledge is not infallible, it is possible to produce warranted generalizations about human organizations by applying the norms of scientific research (Phillips & Burbules, 2000). Post-positivist, methods-focused approaches to evaluation are based on the premise that there should be an objective relationship between researchers and those being researched, and that systematic and valid scientific methods should be used to produce justifiable conclusions.

Pragmatic, Use-Oriented Approaches to Evaluation

Alkin and Christie's (2004) second general approach within the cross-disciplinary field of program evaluation focuses not so much on methods, but on *use*—that is, how the knowledge produced through evaluation can be used by key stakeholders in program decisions. This approach was prompted in part by dissatisfaction with methods-focused evaluation research that did not seem to make much of a difference in policymaking or practice (Weiss, 1998). Patton (2008) characterized this approach by referring to it as part of the "utilization turn" in evaluation with an emphasis on "intended uses by intended users." With this approach, the goal is to design evaluations that produce knowledge that can best inform the decisions practitioners and others must make. This model of evaluation is consistent with what Mertens and Wilson (2012, 2019) refer to as the "pragmatic" paradigm in evaluation, which values the impact of evidence as much as the scientific rigor through which the evidence was developed. The use-oriented approach works from a utilitarian stance, assuming that the worth of an evaluation is not simply defined by the rigor of the methods used, but rather by consequences and results—that is, whether an evaluation "works" to support certain kinds of improvements under certain conditions. With pragmatic, use-oriented approaches, no particular research method or measure is necessarily privileged; rather, methods and measures are designed to match purpose and intended use, and evaluators make choices about what to study based on their knowledge and relationships with stakeholders.

Transformative, Equity-Centered Approaches to Evaluation

A third approach within the field of evaluation emphasizes the idea that evaluation is a *constructivist* process (Mertens & Wilson, 2012, 2019) wherein evaluators make judgments about programs by *valuing* particular goals intended to serve the public interest (Alkin & Christie, 2004; Christie & Alkin, 2008, 2013). Building on this work, many evaluation theorists now make a distinction between generally values-centered approaches to evaluation (Alkin & Christie, 2004; Christie & Alkin, 2008, 2013), on one hand, and, on the other hand, explicitly justice- or equity-centered approaches, which are "transformative" (Mertens & Wilson, 2012, 2019; Mertens & Zimmerman, 2015; Thomas & Campbell, 2020).[1] Transformative approaches often utilize dialogic qualitative methods ethically centered in cultural respect, the promotion of human rights, and reciprocity. Here, the idea is that evaluation is inherently a valuing—and

[1] With our three categories, we do not utilize a category labeled, "constructivist, valuing-oriented approaches," which would have followed directly from Alkin and Christie (2004) and Mertens and Wilson (2012, 2019). However, we found that all of the reports in our body of literature that could be described as taking a "valuing-oriented" approach were specifically "transformative" and "equity-centered" rather than broadly constructivist.

political—activity with the potential for political influence and that evaluators should guard against power imbalances by carefully considering whose interests are and are not served by evaluation and whose voices are and are not included (Greene, 2006; House & Howe, 2000). Along related lines, "culturally responsive evaluation" explicitly centers evaluation in culture and cultural competence, emphatically rejecting the idea that evaluation is culture-free (Haugen & Chouinard, 2019; Hood et al., 2015). Culturally responsive evaluation seeks "to bring balance and equity into the evaluation process" (Hood et al., 2015, p. 283) by recognizing unequal resources and by drawing on the lived experiences of those in marginalized groups (Thomas & Campbell, 2020).

It is important to note that efforts to make issues of equity front and center in evaluation reflect an increasingly influential agenda among some scholars and practitioners in the field of program evaluation, a position that is supported by a significant group of funders and philanthropies (Farrow & Morrison, 2019; Wiggins & Sileo, 2020). Along these lines, a framing paper on equitable evaluation argued that "evaluative work should be designed and implemented in a way that is commensurate with the values underlying equity work" (Equitable Evaluation Initative, 2017, p. 8). As a number of evaluation researchers (Andrews et al., 2019b; Gates, 2017, 2020; Schwandt & Gates, 2016, 2021) have suggested, equity-centered approaches to evaluation raise normative questions about objectivity, methods of evaluation, rigor, the roles of evaluators as agents of change, and professional responsibility. These questions are definitely *not* settled in the field of evaluation (Datta, 2011). To the contrary, these questions and their entanglement with highly politicized issues related to racism and racial justice are currently a point of intense contention within the evaluation field (Hall, 2018).

Researchers' Perspectives and Positionality

As co-authors of this paper, we have substantial histories in the field of teacher education. The first author is a well-known university-based teacher education scholar and practitioner who has written about justice and equity in teacher education for more than 35 years and who has studied accountability initiatives in teacher education over the past 20. The second author has worked on issues related to social justice-oriented policy and practice in teacher education for the past decade. Like some of the scholars reviewed above, we work from the assumption that no approach to teacher preparation evaluation is objective, no approach is apolitical, and no approach is innocent of questions about whose interests are served or undermined, whose perspectives are represented or omitted, and whose voices are included or excluded by the processes and results of particular evaluations.

Fully recognizing that values are inherent in any approach to teacher preparation evaluation, however, we do not take a relative stance in this paper by simply describing variations in recommendations. Rather, we aim to take a stand, following Greene (2006) and others (Farrow & Morrison, 2019; Gates, 2020; Haugen & Chouinard, 2019; House & Howe, 2000; Mertens & Zimmerman, 2015; Schwandt & Gates, 2016, 2021; Thomas & Campbell, 2020), who have argued both that the most defensible values in evaluation are those related to justice, equity, and empowerment and that it is critical to understand how power is taken up in the practice of evaluation. In particular, in this review, we raise questions about the presence, absence, and meanings of equity

and justice in teacher preparation evaluation. Our analysis is grounded in the premise that the work and lives of students, teachers, teacher educators, community members, and evaluators are mediated by long-standing and intersecting systems of inequality (Lather, 1992; Walby, 2007).

REVIEWING TEACHER PREPARATION EVALUATION, 2010-2020

In addition to variations in their conclusions and recommendations, the 19 reports we reviewed differed in format, length, organization, scope, audience, sponsoring agencies, and the larger policy or political agendas to which they were attached. To synthesize and critique these reports, we first grouped them into the three categories introduced in the previous section, based on their underlying assumptions and theories of action related to evaluation. In Figure 1, within each of three categories, the reports are organized chronologically and by organization or lead author.

Post-Positivist, Methods-Focused Approaches to Teacher Preparation Evaluation

As noted in the introduction, the 2010 NRC report on the degree to which teacher preparation was grounded in scientific research was seminal. Although the report cautioned that causal evidence directly linking preparation to student achievement was exceedingly difficult to establish, it also asserted the great need for comprehensive data collection to support quality control and accountability. Many of the reports about teacher preparation evaluation over the next decade, particularly those we have placed in the first and second categories, can be understood, at least in part, as responses to the NRC's call and to the broader policy and political milieu out of which it emerged

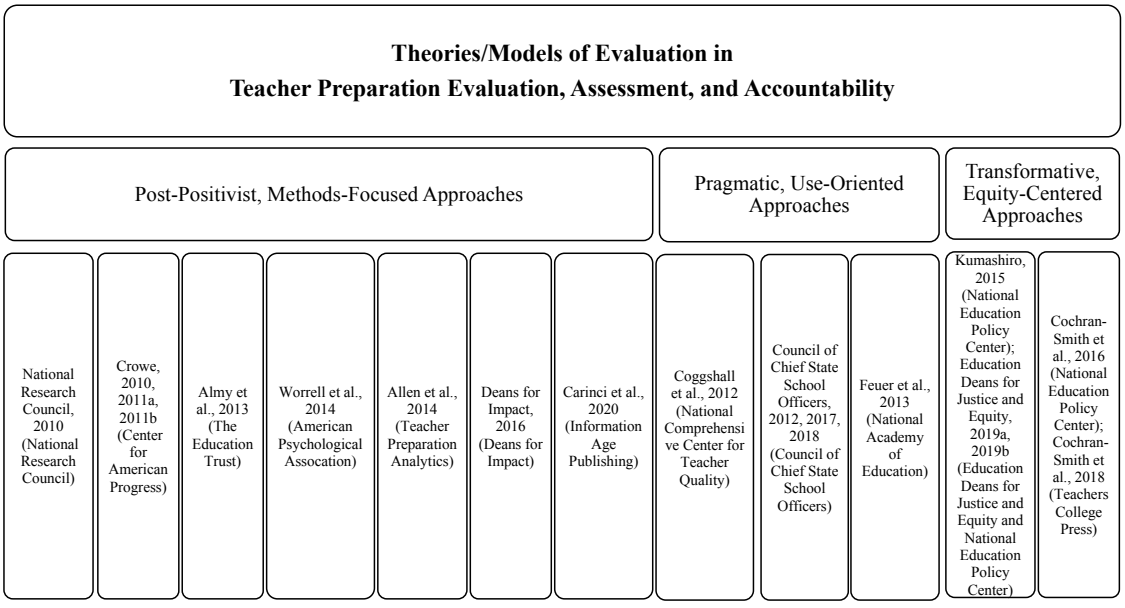


FIGURE 1 Reports on teacher preparation evaluation: Underlying theories/models of evaluation.

TABLE 2 Post-Positivist, Methods-Focused Approaches to Teacher Preparation Evaluation, Assessment, and Accountability, Arranged Chronologically and by Author/Group

Report	Sponsoring/Publishing Organization and Funding Sources	Format and Intended Audience	Purpose of Report	Recommendations for "Best Practices"
National Research Council. (2010). *Preparing teachers: Building evidence for sound policy*. The National Academies Press.	Sponsoring Organization: National Research Council, the operating arm of the National Academies that provides objective policy advice Funding Sources: Carnegie Corporation of New York, Congress, Ewing Marion Kauffman Foundation, Spencer Foundation, U.S. Department of Education	Format: Congressionally mandated 200+ page report Intended Audience: Policymakers	• Synthesize the research base on teacher preparation for policymakers • Propose a research agenda for teacher preparation	• Research agenda that includes improved understanding of the relationships between characteristics of teacher preparation and student learning • Comprehensive, coherent system for collecting data about teacher preparation
Crowe, E. (2010). *Measuring what matters: A stronger accountability model for teacher education*. Center for American Progress. Crowe, E. (2011a). *Race to the Top and teacher preparation analyzing state strategies for ensuring real accountability and fostering program innovation.* Center for American Progress. Crowe, E. (2011b). *Getting better at teacher preparation and state accountability strategies, innovations, and challenges under the federal Race to the Top program.* Center for American Progress.	Sponsoring Organization: Center for American Progress, an independent nonpartisan policy institute Funding Sources: Bill & Melinda Gates Foundation; The Eli and Edythe Broad Foundation	Format: 44-page policy paper/white paper; 42-page policy paper/white paper; 67-page policy paper/white paper Intended Audience: Policymakers	• Offer a redesigned accountability system for teacher preparation with direct regulatory oversight • Describe and analyze one component of the Race to the Top proposals and findings from Race to the Top funded states	• Hold programs accountable for graduates' K-12 pupil achievement • Implement standardized observational assessments • Develop state data systems • Implement feedback surveys of program graduates and employers • Monitor state performance
Almy, S., Tooley, M., & Hall, D. (2013). *Preparing and advancing teachers and school leaders: A new approach for federal policy.* The Education Trust.	Sponsoring Organization: The Education Trust, a national nonprofit that promotes policies to ensure high quality education for all students Funding Source: The Education Trust	Format: 16-page policy brief Intended Audience: Policymakers	• Offer recommendations for federal involvement in teacher preparation improvement and accountability	• Federal government involvement in teacher preparation performance measurement and accountability • Allocate federal competitive funds and other resources

continued

Report	Sponsoring/Publishing Organization and Funding Sources	Format and Intended Audience	Purpose of Report	Recommendations for "Best Practices"
Allen, M., Coble, C., & Crowe, E. (2014). *Building an evidence-based system for teacher preparation.* Teacher Preparation Analytics.	Sponsoring Organization: Council for the Accreditation of Educator Preparation (CAEP) Funding Sources: CAEP, Council of Chief State School Officers, Pearson, Inc.	Format: 155-page report Intended Audience: Policymakers, teacher education programs, national accrediting organizations, educational researchers	• Offer a uniform framework for assessment of teacher preparation program performance • Support the development of a solid performance evaluation system	• Implement 13 key effectiveness indicators across 4 categories of assessment, including candidate selection profile, knowledge and skills for teaching, performance as classroom teachers, and program productivity and alignment to state needs
Worrell, F., Brabeck, M., Dwyer, C., Geisinger, K., Marx, R., Noell, G., & Pianta, R. (2014). *Assessing and evaluating teacher preparation programs*. American Psychological Association.	Sponsoring Organization: American Psychological Association (APA), a leading scientific and psychological professional organization in the United States Funding Sources: APA Board of Educational Affairs, CAEP	Format: 45-page task report Intended Audience: Policymakers, teacher education practitioners	• Use data and scientific methods to evaluate the effectiveness of teacher preparation programs	• Thirteen recommendations that lead to the best use of data for program accountability and improvement, including K-12 student outcome measures, standardized observation protocols, and stakeholder surveys
Deans for Impact. (2016). *From chaos to coherence: A policy agenda for accessing and using outcomes data in educator preparation*. http://www.deansforimpact.org/wp-content/uploads/2016/11/From_Chaos_to_Coherence.pdf.	Sponsoring Organization: Deans for Impact, a national nonprofit organization that brings together leaders in educator preparation Funding Sources: Charles and Lynn Schusterman Family Philanthropies	Format: 20-page policy brief Intended Audience: Deans/leaders in teacher preparation, policymakers	• Offer a policy agenda for outcomes-based accountability and data-informed improvement in teacher preparation	• Improve access to data through policies that provide data on teacher preparation program graduates • Outcomes-focused certification processes
Carinci, J. E., Meyer, S. J., & Jackson, C. (Eds.). (2020). *Linking teacher preparation program design and implementation to outcomes for teacher and students.* Information Age Publishing.	Publishing Organization: Information Age Publishing Funding Source: N/A	Format: 200+ page edited volume Intended Audience: Teacher education programs, national accreditation organization, researchers	• Improve the research base on how to prepare effective teachers through evidence on program structures, policies, and practices	• Improve the research base on how to prepare effective teachers through evidence on program structures, policies, and practices

(Cochran-Smith, 2005; Mehta, 2013). Table 2 provides basic information about the reports in the first group.

Context

In many countries over the past two decades, accountability has come to be regarded as a powerful policy tool for improving teacher preparation, as evidenced in new standards, monitoring systems, and/or auditing procedures for the colleges and universities that prepare teachers. In the United States, the logic of the accountability approach, which is reflected in the reports in this first group, is captured in this string of claims: holding teacher education accountable boosts the quality of preparation programs and institutions; boosting the quality of teacher preparation increases the overall level of teacher quality, especially in terms of students' achievement; and higher levels of student achievement ensure both the prosperity of individuals and the long-term economic health of the nation (Cochran-Smith et al., 2018). The key accountability assumption here is that the way to "fix" teacher education is rigorous public evaluation and monitoring of the inputs, processes, and especially the outcomes of preparation programs, including the impact of their graduates.

Purpose and Values

Most of the reports in this first category were produced in the midst of contentious debates about how teacher preparation should be held accountable to the public, the profession, and consumers. These reports aimed to make evidence-based policy and practice recommendations regarding the improvement of—or in some cases, the complete overhauling of—the state, federal, and/or professional evaluation systems that governed teacher preparation. By labeling the theory of evaluation underlying the reports in this group "post-positivist, methods-focused approaches to teacher education evaluation," we emphasize that these reports zeroed in on the metrics they claimed should be components of all evaluation systems. Central to this approach are the core principles of objectivity and rigor, along with the belief that it is the responsibility of those making policy and practice decisions to utilize the results of evaluation approaches backed by evidence. Underlying these reports is the assumption that the quality of teacher preparation can only be improved when programs are held accountable for outcomes with severe sanctions for programs that do not perform. It is important to note here that although the reports in the post-positivist category adhere to the principles or rigor and objectivity, they are not devoid of values. The key underlying value here is that teacher quality—defined in terms of a uniformly effective work force—should be provided to all students in all schools in the nation, which would presumably remedy current disparities in achievement.

In the early part of the 2010s, the Center for American Progress published a trio of reports by Crowe (2010, 2011a, 2011b), an independent advisor on teacher quality policy and preparation initiatives for multiple public and private agencies, which outlined a federal model for creating "stronger" and "real" accountability in teacher preparation. Crowe asserted that every state should have an evaluation system with four assessments: measures of teacher effectiveness (e.g., value-added assessments linking teacher preparation data to teacher and student achievement data), feedback

from graduates and employers, tests of teacher knowledge and skill that predict performance, and accurate measures of teacher retention. Two years later, The Education Trust, a national nonprofit advocacy organization promoting academic achievement for all students, released a report (Almy et al., 2013) recommending that, under the threat of removal of eligibility for federal funds, the HEA should require all states to hold all preparation programs accountable for the performance of teachers using statewide measures of teacher impact on student growth as well as employment, retention, and program selectivity data.

In 2014, Teacher Preparation Analytics (TPA), a company founded in 2012 to develop high-leverage strategies to strengthen teacher preparation, released a framework for analyzing assessment in teacher preparation and moving toward a more evidence-based system (Allen et al., 2014). The TPA report, which was commissioned by CAEP, proposed a set of "key effectiveness indicators" regarding candidate selection, knowledge and skills for teaching, classroom performance, and program alignment with state needs, to be in place by 2020. Along related but different lines, the report of the American Psychological Association (APA) Task Force on Assessing and Evaluating Teacher Preparation Programs (Worrell et al., 2014), comprised primarily of psychologists affiliated with schools of education, provided empirical support for several of CAEP's controversial standards. The APA report asserted that program assessments at all decision points from entry to post-graduation should be valid and reliable, thus allowing users to make comparisons on an "even playing field." The 2016 report of Deans for Impact, an organization of school of education deans committed to supporting the teacher effectiveness agenda in teacher preparation, advocated for state-level evaluation systems that produced uniform "actionable data" (p. 2) regarding the impact of educator preparation. The dean's group argued that it was precisely the *lack* of "valid, reliable, timely, and comparable data about the effectiveness of the teachers and school leaders they prepare" (p. 2) that had plagued preparation programs for years and prevented them from moving from "chaos" to "data coherency" (p. 3).

Finally, Carinci et al. (2020), education researchers interested in accountability practice and policy, edited the second volume in Information Age Publishing's book series on contemporary issues in accreditation, assessment, and program evaluation research in educator preparation. The volume focused on data-driven accountability linking program design to teacher and student outcomes. This volume was produced after the controversies about Title II HEA regulations and thus, unlike the other reports in this category, was not intended to influence those debates. Nevertheless, the report was intended to inform teacher preparation policy and to drive continuous improvements in practice by using empirical research to open the "black box" between teacher preparation and outcomes for students and teachers.

"Best Practices"

The reports in this first category were grounded in the explicit assumption that the validity of assessment instruments along with their uniform implementation in state-level and/or professional accreditation evaluation systems were the key to boosting the caliber of teacher preparation programs and the teachers whom those programs produced. The reports conceptualized "best practices" in teacher education evalua-

tion in two ways: (1) endorsement of particular methods or statistical approaches to evaluation based on scientific evidence; and (2) identification of the characteristics of exemplary state teacher education evaluation systems.

With the exception of the Carinci volume, all of the reports in this group emphasized that the singular most important aspect of teacher preparation evaluation—arguably, the "best" of the "best practices"—was the statewide use of valid measures of the impact of new teachers on student learning that was also linked to information about the programs that prepared those teachers. The reports issued by The Education Trust (2013) and Deans for Impact (2016) did not specify which measures should determine teacher effectiveness. However, Crowe (2010, 2011a, 2011b) and both the TPA (Allen et al., 2014) and the APA (Worrell et al., 2014) reports recommended value-added methods, suggesting that the considerable problems involved in using these for the evaluation of teachers and teacher preparation programs could and should be overcome. This conclusion has not been supported by measurement experts (American Education Research Association, 2015; Braun, 2005; Easton, 2008; Economic Policy Institute, 2010; Haertel, 2013). In addition, researchers who have studied state systems linking program graduates' value-added teaching scores to their preparation programs as a way to evaluate and improve programs have generally concluded that these systems do not provide information about why results occurred and thus provide little information about how programs might improve (Goldhaber, 2013; Plecki et al., 2012).

Other "best" teacher effectiveness measures recommended by the reports in this first category included standardized protocols for observing classroom practices and teacher interactions assumed to have a direct impact on student learning. Along these lines, the APA report (Worrell et al., 2014) specified the Classroom Assessment Scoring System (CLASS) observation instrument (Pianta & Hamre, 2009) and the observation protocols identified by the Measures of Effective Teaching (MET) Project (2012).

Along with teacher effectiveness measures, the reports in this first group also called for the widespread use of satisfaction surveys of teacher candidates and graduates, and their students and/or employers (Allen et al., 2014; Almy et al., 2013; Crowe, 2010; Deans for Impact, 2016; Worrell et al., 2014); employment, retention, and career trajectory data (Allen et al., 2014; Almy et al., 2012; Crowe, 2010; Deans for Impact, 2016); teacher tests that predict effectiveness (Crowe, 2010); and data regarding program entry selectivity criteria (Allen et al., 2014; Almy et al., 2013), although the APA report (Worrell et al., 2014) specifically concluded that the evidence regarding selectivity criteria was too weak to recommend its widespread use.

In addition to recommending particular approaches and tools, several of the reports made sweeping recommendations regarding "best practices" for state evaluation systems. As noted above, Crowe (2010, 2011a, 2011b) proclaimed that all states should adopt new accountability systems to improve outcomes, including teacher effectiveness measures, standardized observations, retention rates, feedback from graduates and employers, and teacher tests. Crowe pointed to Tennessee and Delaware (first round Race to the Top fund recipients) and to Florida and Louisiana (second and third round Race to the Top fund recipients, respectively) as showing promise. Based on a landscape analysis of the data available in the states of their own members, Deans for Impact (2016) concluded that there was a severe paucity of valid data in most states; they called for a two-pronged policy agenda to provide the necessary data about teacher

effectiveness in all states and create a new outcomes-focused certification process that elevated effectiveness-centered programs. The Education Trust report (Almy et al., 2013) explicitly called for redesigned HEA state reporting requirements with performance measures tied to federal funding and other resources. The TPA report (Allen et al., 2014) reviewed 15 sample states according to the report's proposed key effectiveness indicators, concluding that although some states had made progress, implementing these indicators was beyond the current efforts of the states.

Although different in many ways, the reports in this first group were remarkably consistent in purpose and specific recommendations regarding teacher preparation evaluation. They emphasized externally-driven, outcomes-based evaluation systems featuring valid assessments of the impact of new teachers on student learning, standardized observation protocols of classroom practices, and interactions with a direct impact on student learning, satisfaction surveys, and employment information. They assumed that implementing comprehensive external systems of evaluation with the components specified above and making the data from these systems publicly available would not only hold teacher preparation accountable but would also dramatically improve teacher preparation by identifying strong programs and forcing weak programs to improve or exit the field.

Pragmatic, Use-Oriented Approaches to Teacher Preparation Evaluation

As we demonstrate below, the reports in the pragmatic, use-oriented category draw on many of the same purposes, values, and assumptions as the reports analyzed in the previous post-positivist, methods-focused group. However, the reports in the pragmatic category differ from the prior category in that they prioritize usability of evaluation findings by intended users, key decision makers, or program leaders (Alkin & Christie, 2004; Mertens & Wilson, 2012, 2019), as well as alignment among evaluation purposes, use, selection of tools, audience, and weighing the strengths and weaknesses of various assessments. Furthermore, the reports in the pragmatic, use-oriented category emphasize the use of trustworthy evidence that is of interest to specific audiences (e.g., policymakers, professional organizations, teacher preparation programs). Table 3 lists the reports in this second group.

Context

Almost all of the reports in the pragmatic, use-oriented category were published in the first half of the 2010s in response to debates about federal and state regulations, the development of new professional standards, and the growing number of philanthropies and independent advocacy organizations involved in teacher preparation evaluation. In particular, some of the reports in this category responded explicitly to the Obama administration's proposed reform plan, *Our Future, Our Teachers*, to proposed revisions to Title II HEA reporting regulations, and to Race to the Top funding requirements. The reports published in the latter half of the decade reflected continued efforts to shape teacher preparation evaluation through projects such as the CCSSO NTEP.

TABLE 3 Pragmatic, Use-Oriented Approaches to Teacher Preparation Evaluation, Assessment, and Accountability, Arranged Chronologically and by Author/Group

Report	Sponsoring/Publishing Organization and Funding Sources	Format and Intended Audience	Purpose of Report	Recommendations for "Best Practices"
Coggshall, J. D., Bivona, L., & Reschly, D. J. (2012). *Evaluating the effectiveness of teacher preparation programs for support and accountability.* National Comprehensive Center for Teacher Quality.	Sponsoring Organization: National Comprehensive Center for Teacher Quality, a collaborative effort among Educational Testing Service, Learning Point Associates/ American Institutes for Research, and Vanderbilt University Funding Source: U.S Department of Education	Format: 52-page research and policy brief Intended Audience: State-level policymakers	• Provide a resource for state education agency personnel and other state-level stakeholders to use as they redesign systems of teacher preparation program accountability and support	• Multiple measures exist for assessing the quality of teacher preparation programs; newer measures hold promise for providing a greater understanding of the quality of teacher preparation programs • Bridge the divide between current data and evaluation capacity and what is needed for accountability, continuous program improvement, and equity
Council of Chief State School Officers. (2012). *Our responsibility, our promise: Transforming educator preparation and entry into the profession.* Task Force on Educator Preparation and Entry into the Profession. Council of Chief State School Officers. (2017). *Transforming educator preparation: Lessons. learned from leading states.* Council of Chief State School Officers (2018). *Measuring what matters: Recommendations from states in the network for transforming educator preparation.*	Sponsoring Organization: Council of Chief State School Officers (CCSSO), a nonpartisan, nationwide, nonprofit organization of public officials who head state departments of elementary and secondary education across the United States Funding Sources: Bill & Melinda Gates Foundation, CCSSO, Charles and Lynn Schusterman Family Philanthropies, The Joyce Foundation, S. D. Bechtel, Jr. Foundation	Format: 47-page policy report; 24-page follow-up report Intended Audience: State-level policymakers and leaders, teacher preparation programs	• Transform teacher preparation through licensure, program approval, and data collection, analysis, and reporting • Offer a "playbook" for states' efforts to strengthen teacher preparation	• States will hold preparation programs accountable by exercising authority to determine which programs should operate and recommend candidates for licensure, including establishing a clear and fair performance rating system to guide continuous improvement • Raising the bar for teachers through licensure reform • Rigorous preparation program standards, evaluation, and approval • Use data to measure success and continuous improvement

continued

Report	Sponsoring/Publishing Organization and Funding Sources	Format and Intended Audience	Purpose of Report	Recommendations for "Best Practices"
Feuer, M. J., Floden, R. E., Chudowsky, N., & Ahn, J. (2013). *Evaluation of teacher preparation programs: Purposes, methods, and policy options.* National Academy of Education.	Sponsoring Organization: National Academy of Education, a nonprofit, nongovernmental organization dedicated to advancing high-quality research Funding Source: National Science Foundation	Format: 130-page report Intended Audience: Federal government, states, media and independent organizations, nongovernmental organizations, teacher preparation programs	• Explain variation across teacher preparation evaluation systems • Provide guidance in the development of new and better evaluation systems	• Develop a coherent evaluation system that serves its intended purposes and leads to valid inferences for accountability, consumer protection, and program improvement • Design coherent evaluation systems across purpose, use, methods, and tools, consistent with the values of program leaders and the intended users of the evaluation • Consider sources of evidence, inferences, and incentives

Purposes and Values

Like the reports in the first category, the reports in the pragmatic, use-oriented category offered evidence-based recommendations intended to contribute to "new and better" systems of teacher preparation evaluation, accountability, and assessment. Exploring multiple measures and sources of evidence, these reports described how various stakeholders—including the federal government, state policymakers, national programmatic accrediting agencies, independent media and non-governmental organizations, and teacher preparation programs—could utilize preparation program evaluation to improve teacher preparation, teacher quality, and K-12 student learning. These reports built on the core principle of validity and offered frameworks and promising examples of program evaluation, accountability, and/or assessment systems. These reports also zeroed in on *alignment* across intended purpose, use, values, audience, measures, and stakeholders as a key aspect of evaluation done well. In this category, the reports assumed that evaluation should be tied to the interests and values of the audience.

For example, in a research and policy brief released by the National Comprehensive Center for Teacher Quality (NCCTQ, not to be confused with NCTQ), a collaborative effort between the Education Commission of the States, Education Testing Service, Learning Points Associates, and Vanderbilt University, Coggshall et al. (2012) called for "rethinking" teacher preparation accountability through a more "results-oriented approach" (p. 2) to program evaluation at the state level. The NCCTQ brief aimed to provide a resource for state education agencies to develop evaluation systems through an analysis of measures associated with program processes and outcomes, arguing: "the success and usefulness of [teacher preparation] accountability efforts are dependent on the quality of the measures and how states, teacher preparation programs, and individuals use the data gathered from these measures" (p. 5). Framed by concepts of reliability, validity, and best use, the NCCTQ report called for additional "research and capacity building … to bridge the divide between current data and evaluation capacity, and what is needed for accountability, program improvement, and equity" (p. 34). Even though equity was mentioned as a purpose of state evaluation systems, NCCTQ's attention to equity was limited to identifying programs that prepared high quality teachers for "high-need schools" and "traditionally underserved populations," with the assumption that the redistribution of teachers would address disparities in schooling and society.

At about the same time, the CCSSO Task Force on Educator Preparation and Entry into the Profession (2012) published a report offering guidance to state education agencies and policymakers with the intention of transforming the profession through state policy levers, such as teacher licensure, program approval, and data collection, analysis, and reporting. The CCSSO Task Force, composed of state education leaders and policymakers, aimed to support program accountability and continuous improvement through rigorous and transparent standards and rating systems aimed to ensure that "learner-centered" teachers could support K-12 students in meeting college- and career-ready standards. To do so, the Task Force recommended that program approval standards be met in a variety of ways using multiple measures. Toward the end of the decade, CCSSO (2017, 2018) published two follow-up reports, highlighting "leading state efforts" to transform educator preparation through NTEP. The follow-up reports

described lessons learned from states' efforts to transform teacher preparation, including "early challenges" and "essential actions for achieving results" (CCSSO, 2017, p. 6) primarily through collaboration among state policymakers, agencies, districts, schools, and teacher educators.

Addressing a research, policy, and teacher education practitioner audience, the NAEd report *Evaluation of Teacher Preparation Programs: Purposes, Methods, and Policy Options* (Feuer et al., 2013) was guided by a steering committee of deans of schools of education, faculty, and policymakers. The NAEd report was intended to clarify the many variations in teacher preparation evaluation systems, with the underlying assumption that evaluation was a "necessary ingredient" to improving teaching and learning. The NAEd report acknowledged the multiple purposes of evaluation, including accountability, consumer information and protection, and programmatic improvement, and it sorted out the many entities involved in evaluation. The report also analyzed the strengths and limitations of various sources of evidence in evaluation systems, arguing that any system, set of measures, or source of evidence should be based on principles of validity that lead to defensible conclusions so that various entities could use the results of evaluation to make sound decisions.

"Best Practices"

Like the reports in the prior post-positivist category, the reports in this pragmatic category were guided by validity as the primary criterion for assessing preparation program evaluation systems. However, the reports in the pragmatic category explicitly attended to alignment across teacher preparation program evaluation systems for multiple purposes. In the NCCTQ (Coggshall et al., 2012) and CCSSO reports (CCSSO, 2012, 2017, 2018), the authors called for state-level evaluation systems as key levers for improving teacher preparation, teaching, and learning while recognizing the multiple stakeholders involved in the process of designing, implementing, and using these systems. The NAEd report (Feuer et al., 2013) analyzed purposes, methods, and policy options in teacher preparation program evaluation involving multiple organizations and agencies. Overall, these reports conceptualized "best practices" in teacher preparation program evaluation in terms of (1) alignment across evaluation purposes, measures, and use, and (2) engagement and use by multiple stakeholders and groups.

In terms of alignment, the NAEd report proposed a list of guiding questions for developing teacher preparation evaluation systems, starting with identifying the purpose of the evaluation, articulating the aspects of teacher preparation that "matter most," determining the sources of evidence that would yield accurate and useful information, and monitoring the evaluation system for intended and unintended consequences. Along similar lines, the NCCTQ report recommended that "states and other organizations, in collaboration with stakeholder groups, should consider the strengths and the weaknesses of the available measures and select those that will best fit the context of the evaluation" (p. 34).

Unlike the post-positivist reports, the pragmatic reports did not advocate for specific measures. Rather, the reports in this group analyzed the strengths and weaknesses of multiple measures and cautioned against any one measure as the sole or primary source of evidence in an evaluation system. Specifically, the reports in this group ana-

lyzed input/process measures including teacher candidate selection criteria (CCSSO, 2012, 2017, 2018; Coggshall et al., 2012; Feuer et al., 2013), course syllabi (Coggshall et al., 2012; Feuer et al., 2013), and faculty qualifications (Feuer et al., 2013). They also identified common measures tied to clinical experience, such as quality or number of hours of clinical experience (CCSSO, 2012, 2017; Coggshall et al., 2012; Feuer et al., 2013).

Firmly grounded in an outcomes-based approach to evaluation, the reports in the pragmatic category also analyzed many of the same output/outcomes measures as did the reports in the first group, including licensure tests and performance assessments of teacher candidate knowledge and skills (CCSSO, 2012, 2017; Coggshall et al., 2012; Feuer et al., 2013); K-12 student achievement, including growth modeling and value-added models (CCSSO, 2017, 2018; Coggshall et al., 2012; Feuer et al., 2013); teacher evaluation and classroom observations (CCSSO, 2017, 2018; Coggshall et al., 2012; Feuer et al., 2013); employer surveys (Coggshall et al., 2012; Feuer et al., 2013); program graduate surveys (Coggshall et al., 2012; Feuer et al., 2013); hiring and placement data (Coggshall et al., 2012); and retention data (Coggshall et al., 2012). It is important to note that some of these reports outlined questions and limitations of specific measures, especially value-added approaches and growth modeling as a method for evaluating preparation programs, citing the technical and logistical challenges associated with these measures and calling for further research on their use (Coggshall, 2012; Feuer et al., 2013).

Together, the reports in the pragmatic category sorted out many actors and organizations involved in teacher preparation program evaluation in terms of the purposes of accountability, consumer protection, and programmatic improvement. These included federal and state education agencies (CCSSO, 2012, 2017; Coggshall et al., 2012; Feuer et al., 2013) along with non-governmental, independent organizations, media, national programmatic accrediting agencies, and teacher preparation programs. For example, in the CCSSO (2017) follow-up report, state exemplars were those that featured strong collaboration and engagement with state education agencies, district superintendents, school administrators and teachers, and preparation programs housed in institutions. Overall, the reports in this category recognized the multiple purposes and stakeholders involved in teacher preparation evaluation and aimed to enhance the usefulness of evaluation systems based on the purposes, audience, and values of the stakeholders and audiences who would use the resulting evaluation.

Transformative, Equity-Centered Approaches to Teacher Preparation Program Evaluation

As we describe below, the reports in this category differ significantly from those in the previous two categories in terms of purposes, values, and assumptions. They fit within a transformative paradigm of evaluation, prioritizing evaluations intended to serve public purposes, such as democracy, equity, and justice, and explicitly addressing issues of power and privilege (Mertens & Wilson, 2019). Table 4 lists the reports in the third group.

TABLE 4 Transformative, Equity-Centered Approaches to Teacher Preparation Evaluation, Assessment, and Accountability, Arranged Chronologically and by Group

Report	Sponsoring/Publishing Organization and Funding Sources	Format and Intended Audience	Purpose of Report	Recommendations
Kumashiro, K. K. (2015). *Review of proposed 2015 federal teacher preparation regulations*. National Education Policy Center.	Publishing Organization: National Education Policy Center, a policy center housed at the University of Colorado Boulder, committed to producing and disseminating high-quality, peer-reviewed research to inform education policy decisions Funding Source: N/A	Format: 15-page brief Intended Audience: Researchers, teacher educators, policymakers	• Analyze the evidentiary support for the proposed 2015 federal teacher preparation regulations • Identify concerns about to the regulations	• Regulations should consider the added burden to institutions • Regulations should consider the impact of a teacher in relation to larger systems • Regulations should minimize barriers to the profession, particularly for under-represented groups • Regulations should consider a broadly the goals of publication education
Education Deans for Justice and Equity. (2019a). *Seven trends in U.S. teacher education and the need to address systemic injustices*. National Education Policy Center. Education Deans for Justice and Equity. (2019). *A framework for assessment and transformation* ("The EDJE Framework" v.6.20.2019).	Publishing Organization: National Education Policy Center, a policy center housed at the University of Colorado Boulder committed to producing and disseminating high-quality, peer-reviewed research to inform education policy decisions Funding Source: N/A	Format: 16-page report; 21-page framework Intended Audience: Researchers, teacher educators, policymakers	• Analyze current trends in teacher preparation, including federal regulations, national programmatic accreditation, state level policies, and the role of advocacy groups • Offer a framework for assessment and transformation in teacher education/ schools of education for action planning and implementation	• Framework for institutional assessment and strategic planning • Assumes educational institutions are not politically or ideologically neutral • Considers the role of schooling in relation to a larger society • To advance justice and equity in education must consider the systems broadly in the following categories: (a) governance and finances, (b) teaching and learning, (c) faculty and staff, and (d) partnerships and public impact

continued

TABLE 4 Continued

Report	Sponsoring/Publishing Organization and Funding Sources	Format and Intended Audience	Purpose of Report	Recommendations
Cochran-Smith, M., Stern, R., Sánchez, J. G., Miller, A., Keefe, E. S., Fernández, M. B., Chang, W. C., Carney, M. C., Burton, S., & Baker, M. (2016). *Holding teacher preparation accountable: A review of claims and evidence.* National Education Policy Center.	Publishing Organization: National Education Policy Center, a policy center housed at the University of Colorado Boulder committed to producing and disseminating high-quality, peer-reviewed research to inform education policy decisions	Format: 34-page brief; 228-page book Intended audience: Researchers, teacher educators, policymakers	• Analyze major national initiatives intended to improve teacher quality by "holding teacher education accountable" for its arrangements and/or its outcomes • Examine the claims and evidentiary support behind the dominant accountability paradigm • Propose an alternative for the dominant accountability paradigm	• Policymakers must acknowledge multiple factors that that influence student outcomes, including teacher quality • Evaluation systems should support programmatic improvement and change • Offer a framework for democratic accountability in teacher education across eight dimensions with themes of strong democracy, strong equity, democratic accountability as an alternative to the dominant paradigm, intelligent professional responsibility, and deliberative and critical democratic education
Cochran-Smith, M., Carney, M. C., Keefe, E. S., Burton, S., Chang, W.C., Fernandez, M. B., Miller, A. F., Sanchez, J. G., & Baker, M. (2018). *Reclaiming accountability in teacher education.* Teachers College Press.	Publishing Organization: Teachers College Press Funding Source: N/A			

Context

The reports we placed in the transformative, equity-centered category were published in the latter half of the 2010s. This time period was during and immediately following the highly contentious debates about federal and state regulations, national programmatic accreditation standards, report cards published by independent advocacy agencies, and specific measures aimed at teacher preparation evaluation and accountability.

Purposes and Values

The reports in the transformative, equity-centered category were written by teacher education researchers, practitioners, and leaders of schools of education with expressed commitments to justice and equity. Underlying these reports were three key assumptions. First, the reports assumed that teacher preparation evaluation is fundamentally value-laden, inherently political, and attached to broader agendas. Thus, the reports did not aim to be "objective" in terms of approaches, measures, purposes, and consequences of teacher preparation program evaluation. Second, teacher preparation and teacher quality were regarded as part of larger policy and political systems, not independent factors in educational success. Third, teacher preparation program evaluation was considered in relation to broader equity and democratic projects wherein education was viewed as a public enterprise for the common good, with the aims of facilitating deliberative and critical discourse and democratizing knowledge and participation. From this lens, a key purpose of education is to challenge inequities for students, families, and communities through strong or transformative democracy.

Together, the reports in this category critiqued the major teacher preparation evaluation and accountability initiatives of the 2010s, unpacking their underlying assumptions and assessing their evidentiary support. In addition, these reports sought to reframe and offer alternatives to the "common sense" and market-based discourses and approaches to evaluation that had become dominant since the late 1990s. In a policy brief published by the National Education Policy Center (NEPC), Kumashiro (2015), a former education dean and founding member of Education Deans for Justice and Equity (EDJE), reviewed the proposed and highly controversial federal teacher preparation reporting regulations under Title II of the HEA, explicating specific concerns and critiquing the lack of inclusive, democratic decision-making around the regulations.

At about the same time, NEPC also released a policy brief by a group of teacher educators led by Cochran-Smith (Cochran-Smith et al., 2016), which analyzed four key teacher education accountability initiatives—the HEA Title II reporting requirements approved in 2016, CAEP accreditation standards, NCTQ teacher preparation reports cards, and edTPA. Through an analysis of policy claims and evidence, the brief concluded that these accountability initiatives were generally based on "thin evidence"—that is, limited evidentiary support that the policies actually had the capacity to work as levers for teacher preparation improvement—and on "thin equity"—that is, they failed to account for the multiple, complex in- and out-of-school factors in addition to teacher quality that perpetuate inequity for students, families, and minoritized communities (Cochran-Smith et al., 2016). Building on this policy brief, Cochran-Smith et al. (2018) wrote the book, *Reclaiming Accountability in Teacher Education*, which analyzed the

emergence of teacher preparation's accountability era, proposed an eight-dimensional framework for understanding competing accountability policies, and critiqued major national accountability initiatives in terms of these dimensions. The book called for "reclaiming" accountability in teacher education based on "strong equity" and "intelligent professional responsibility," concepts that we return to below.

Building on the work of Kumashiro (2015) and Cochran-Smith et al. (2016, 2018), EDJE (2019a, 2019b), a nationwide alliance of more than 300 current and former leaders of colleges, schools, and departments of education, released a policy brief that outlined problematic trends in teacher preparation including widespread teacher education external accountability mechanisms. The EDJE report (2019a) critiqued these reform efforts for obscuring "the legacies of systemic injustices" and "focusing narrowly on student achievement, teacher accountability, rewards, and punishments" (p. 3). To address these critiques, EDJE (2019a, 2019b) developed a comprehensive Framework for Assessment and Transformation to guide the work of schools of education. EDJE (2019a, 2019b) called for evaluation systems that recognized teacher preparation as part of broader systems that include governance and finance, faculty and staff, teaching and learning, and partnerships and public impact, and that lead to genuine improvement in teacher preparation.

"Best Practices"

In contrast to the reports in the prior two categories, the reports in the transformative category positioned equity and democracy at the center of evaluation. These reports rejected the concept of "best practices" in teacher preparation evaluation, instead proposing accountability and evaluation frameworks that (1) recognize power inequities across multiple individual, institutional, and ideological systems; (2) balance external and internal accountability through democratic processes; and (3) provide information and feedback that lead to improvement in keeping with democratic processes and deliberation and advancing the aims of equity.

The reports in the transformative category called for teacher preparation program evaluation, assessment, or accountability systems that acknowledge the broader systems and structures that perpetuate inequities and injustices. The EDJE (2019a) report argued that "teacher education should be guided by a deep understanding of the roles of schools and universities within a larger society" (p. 6). EDJE offered a framework with 13 priority work areas grouped in 4 categories to identify power structures and systems that perpetuate inequity in order to dismantle them. Along similar lines, in their framework for democratic accountability in teacher preparation, Cochran-Smith et al. (2018) called for accountability and evaluation systems that recognize and challenge systems, structures, and processes that perpetuate inequities as they relate to teaching and teacher education. This involves reframing and expanding measures, approaches, and processes that are part of teacher preparation program evaluation systems.

The reports in the transformative category also argued for systems of teacher preparation evaluation that balance external and internal accountability. Cochran-Smith et al. (2018) proposed the notion of "intelligent professional responsibility" in teacher preparation, linking the concept of "intelligent accountability" (O'Neill, 2013) with the argument that *external* accountability structures should create the conditions for capac-

ity building and collaboration among multiple stakeholders, leading to strong *internal* accountability and professional trust through joint decision-making and participation (Fullan et al., 2015). Along similar lines, the EDJE reports (2019a, 2019b) called for teacher preparation program assessment that is accountable to and works in solidarity with families and communities.

The reports in the transformative, equity-centered category differed from both the reports in the post-positivist, methods-focused category, which called for explicit or single measures, and from the reports in the pragmatic, use-oriented category, which do not privilege or prioritize particular aims of evaluation, but rather acknowledge multiple purposes depending on users. In contrast, the reports in the transformative category work from the explicit aim of centering equity, using multiple measures tailored to local contexts. As Cochran-Smith et al. (2018) argued, this kind of evaluation and accountability "does not assume that all teacher education programs would meet the same goals or use the same assessments, but it does assume that all teacher preparation programs would be responsible for preparing teachers to identify and challenge inequities in school and society and prepare their students to live and work in a democratic society" (p. 169). For the reports in the transformative category, the trustworthiness of evaluation measures and approaches is determined by multiple factors, including the extent to which they address issues of power and privilege and the extent to which they authentically represent the voices of those who have a genuine stake in teacher preparation as part of the evaluation process. Together, the reports in the transformative, equity-centered category aim to reframe teacher preparation evaluation away from the dominant market-oriented, external accountability paradigm and toward an equity-centered, democratic system based on strong equity and intelligent professional responsibility.

CROSS-CUTTING COMMENTS: TEACHER PREPARATION EVALUATION/ASSESSMENT/ACCOUNTABILITY

As our analysis indicates, there is no lack of interest in teacher education evaluation, assessment, and accountability, and many groups have weighed in on this topic. Looking across the three categories we used to organize the reports reviewed in this paper, it is clear that there are important similarities as well as marked differences.

"Best Practices" and the Logic of Accountability

The 14 reports in the post-positivist and pragmatic categories share the general premise that teacher preparation evaluation has three possible purposes: (1) holding programs accountable for program impact, graduates' effectiveness, and outcomes such as retention and program satisfaction; (2) providing trustworthy information to the public and potential consumers about the quality of programs and the teachers they prepare; and/or (3) providing empirical evidence for program improvement, often by identifying and elevating strong programs, while also identifying and exposing weak programs, thus prompting them to improve or discontinue their work as providers of preparation. With most of the reports in the post-positivist and pragmatic categories, the key to all three purposes is assumed to be the use of assessments that yield

valid inferences. For the reports in the post-positivist category, these assessments are generally intended to be coupled with state-level data systems linking program data with data on student achievement, teacher performance, program satisfaction, and/or retention. For the reports in the post-positivist category, "best practice" in teacher education evaluation is thus defined as the widespread implementation of validated, standardized, and uniform measures that assess programs' and graduates' effectiveness in keeping with larger top-down approaches to education reform. The logic of the reports in the pragmatic category is different. Although these reports also identify validity as the key to the selection of measures, they do not define "best practice" in terms of particular metrics and assessment tools, and their recommendations do not necessarily coincide with top-down policy approaches. Rather, the reports in the pragmatic category emphasize how evaluation tools and assessments are used by various audiences, which depends on alignment among evaluation purposes, the selection of appropriate measures, and usability for stakeholders. The logic is thus dependent on the intentions, purposes, and values of the users of evaluation.

The logic of the reports in the transformative category diverges from both those in the post-positivist category and, in a different way, from those in the pragmatic category. As we have shown, the reports in the transformative category intentionally reject the logic and assumptions of the reports in the post-positivist category. In particular, they reject the assumption that implementation of high-stakes, externally-driven accountability systems rooted in a market-based logic will produce substantive and transformative change unless the staggering economic inequities in the nation are addressed (Mathis & Trujillo, 2016). Furthermore, the reports in the transformative category are grounded in research that has indicated that top-down, high-stakes evaluation systems and policies requiring standardized practices in teacher preparation tend to foster superficial compliance, deprofessionalization, and uniformity rather than genuine transformation and attention to local problems (Bell & Youngs, 2011; Johnson et al., 2005; Kornfeld et al., 2007; Valli & Rennert-Ariev, 2002). In contrast, the reports in the transformative category are similar to the reports in the pragmatic category in that they assume that teacher preparation evaluation should be tied to the interests and values of relevant audiences and user groups rather than be predetermined by the designation of particular assessments that are considered to be consistent with a state- or federal-level outcomes-based approach to education reform. However, in contrast to the reports in the pragmatic category, the reports in the transformative category go beyond recognition of multiple stakeholders. They also assert that accountability and evaluation should be reclaimed and reinvented by the profession in collaboration with representatives from the groups most affected by inequities and should be directly guided by principles related to democratic education, justice, and especially strong equity.

Teacher Preparation Evaluation and Equity

The terms "equity" and "justice" appear very few times in the 14 reports that comprise the post-positivist and pragmatic categories, while the terms "accountability," "effectiveness," "data systems," and "validity" appear repeatedly. It would be incorrect, however, as we have noted above, to conclude that there is not an equity aspect to some of these reports. The reports in the post-positivist category assume that one goal

of holding teacher education accountable to the outcomes that the reports recommend is ensuring that "all" of the nation's children and "all" of the nation's schools have access to quality teachers. This perspective on equity rests on two premises—first, that teacher quality/teacher effectiveness is the most important school factor in students' achievement (Darling-Hammond, 2000; Sanders & Horn, 1998), and second, that schools with large numbers of minoritized students and/or students living in poverty conditions are the least likely to have access to effective and fully-qualified teachers (Carter, 2013; Cohen-Vogel & Hunt, 2007). The Education Trust (2008, 2017) refers to the combination of these two premises as "the teacher quality gap" (The Education Trust, 2008, 2017), which is presumed to exacerbate the "achievement gap."

The concept of equity implicit in nearly all of the reports in the first category is based on the assumption that lack of access to teacher quality is a primary *cause* of educational and societal inequity, and thus that redistribution of access to teacher quality is a primary *cure* for inequity. In other words, it is assumed that the redistribution of educational resources, especially teacher quality, has the power to close the gaps that separate minoritized students and students living in poverty conditions from their more economically, politically, and socially advantaged peers. These assumptions are consistent with the larger notion, prominent in social policy since Lyndon Johnson's era, that poverty and income inequality are problems that are "susceptible to correction" through education (Kantor & Lowe, 2013). With regard specifically to teacher preparation evaluation, the reports in the post-positivist category assume that equity is more or less a by-product of a system wherein all school students have teachers whose preparation programs (and candidates) have been evaluated (and met established criteria) according to rigorous, evidence-based, and valid metrics regarding teacher performance, impact, and career trajectories. As noted above, Cochran-Smith et al. (Cochran-Smith & Keefe, in press; Cochran-Smith et al., 2016, 2018) have referred to this perspective as *thin equity* because it assumes that students' equal access to teacher quality—achieved through redistribution—can fix inequity without addressing the systems and structures of power and privilege that produce and reproduce inequity in the first place. Conceptualized from a thin equity lens, evaluation does *not* necessarily account for the need to redistribute education resources, including teacher quality, *as well as* resources well beyond education; recognize differences across learners and communities; and authentically represent the voices of minoritized groups and key stakeholders in deliberations. Evaluation from a thin equity perspective thus tends to mask the structural and systemic barriers that perpetuate inequality, including its racialized nature (Au, 2016).

As we pointed out above, the reports in the pragmatic category concentrate on the use of validated tools and assessment systems that are consistent and aligned with the purposes of intended users and audiences, which means that the reports in this category vary in terms of their attention to equity. Most are intended primarily for supporting state-level policymakers working to redesign their evaluation and accountability systems so that they focus on outcomes (CCSSO, 2012, 2017; Coggshall et al., 2012); thus, their assumptions related to equity are similar to those of the reports in the post-positivist category, described above. The NAEd report (Feuer et al., 2013), however, was intended for multiple audiences with varying evaluation purposes. In this sense,

the NAEd report is not wedded to a particular view of equity because this depends on the values and intentions of its users.

Unlike the reports in the first and second categories, the reports in the third transformative category work from the perspective of strong equity, arguing that issues of justice and equity should be front and center in all aspects of evaluation and accountability (Cochran-Smith et al., 2016, 2018; EDJE, 2019a, 2019b). Drawing on and applying to teacher preparation the social justice theories of political philosopher Nancy Fraser (2003, 2009)[2] and others, Cochran-Smith et al. have defined strong equity in teacher education accountability and evaluation in terms of four dimensions—redistribution (a socioeconomic dimension), recognition (a cultural dimension), representation (a political dimension), and reframing (a discursive dimension). These dimensions draw primarily from Fraser and are elaborated on elsewhere (Cochran-Smith, 2010; Cochran-Smith & Keefe, in press; Cochran-Smith et al., 2016, 2018); we return to these dimensions in the final section below.

"BEST PRACTICES" FOR EVALUATING TEACHER PREPARATION

We were charged with concluding this paper with recommendations regarding "best practices" for evaluating teacher preparation. To do so, we turn once again to ideas from evaluation theory.

The evaluation scholars Schwandt and Gates (2016) argue that evaluation should function in service of an equitable society by providing a "critical voice in social, political, and cultural environments indelibly marked by significant inequalities, power differentials, uncertainty, ambiguity, and interpretability" (p. 67). They suggest that in this sort of environment, evaluation should provide a kind of social conscience:

> Our goal is to push the practice of evaluation further into the domain of a normative undertaking that tackles the questions, "Are we doing the right thing?" and "What makes this the right thing to do?" as opposed to being content with remaining a positive practice largely concerned only with the question of "Are we doing this right?" … [E]valuators need a new way of working that is more directly targeted to the goal of developing evaluation as a critical voice and cultivating the critical voice of others. (pp. 67-68)

Here, we use Schwandt and Gates's (2016, 2021) distinction between "doing things right" and "doing the right thing" to raise questions about "best practices" in teacher preparation evaluation. We have enclosed the phrase "best practices" within scare quotes throughout this paper to signal that we are problematizing this term.

In teacher preparation evaluation, the term "best practices" has the same valence as Schwandt and Gates's question, "Are we doing this right?" In other words, "best practices" is related conceptually to the kinds of instrumental questions that animate many of the reports we reviewed, such as: "Can we overcome the difficulties involved to develop a value-added measure that works to assess teacher preparation program

[2] Cochran-Smith (2010) used Fraser's two dimensions of social justice (redistribution and recognition) in earlier efforts to theorize teacher education for social justice, prior to using an adaptation of Fraser's development of three dimensions in theorizing accountability (Cochran-Smith et al., 2016, 2018). Detailed elaboration of Fraser's ideas, related concepts and literature, and their application to teacher education accountability and evaluation is included elsewhere (Cochran-Smith & Keefe, in press; Cochran-Smith et al., 2016, 2018).

quality in terms of graduates' impact on student achievement?" "Do the observation protocols used by preparation programs provide information for making valid inferences about teacher performance?" "Which evaluation tools align conceptually and methodologically with an organization's or state's evaluation purpose?" "Is there adequate empirical evidence to stipulate teacher candidate selection criteria as part of program evaluation?" and "What kind of state evaluation system most effectively drives outcomes-based accountability in teacher preparation?" Notably, all of these queries focus on technical, methodological, and/or instrumental aspects of getting evaluation "right." Just as notably, none is related to the normative questions, "Are we doing the right thing?" and "What makes this the right thing to do?" Along these lines, the current status of evaluation in teacher preparation seems to mirror the status quo of the field of evaluation more generally, at least as perceived by some groups, including the Equitable Evaluation Initiative (2017), which charged: "evaluation seems to be among the last organizational functions to be examined and revamped through an equity lens." This is not to suggest that "doing things right" is not important, but it is to suggest that this technical question makes sense only within the context of larger ethical questions related to "doing the right thing."

In our analyses in the preceding sections of this paper, we identified the "best practices" recommended for teacher preparation evaluation in the reports located in the post-positivist and pragmatic categories, and we showed that there is a high level of consistency across many of these reports. We also suggested that the concept of "best practices" is not conceptually consistent with the reports located in the transformative category. In keeping with our analyses and critiques throughout this paper, then, we ourselves do *not* propose "best practices" for teacher preparation evaluation as we conclude this paper. Rather, consistent with the normative question "Are we doing the right thing?" we call for a different way of conceptualizing and enacting teacher preparation evaluation that draws on "guiding principles" rather than "best practices." The distinction we are making here is conceptual in that the phrase "guiding principles" is intended to signal that teacher preparation evaluation, which we argue should be strong equity–centered, must be understood as normative, critical, and context-specific. In contrast, the term "best practices" signals "proven" methods in the sense used by some governmental agencies such as the U.S. Department of Education's What Works Clearinghouse, which suggests that teacher preparation evaluation can be objective, uniform, and decontextualized.

In problematizing the notion of "best practices," we are *not* asserting that *none* of the approaches considered as "best practices" in the reports we reviewed or the practices currently in place as part of some state evaluation systems could be part of strong equity–centered evaluation systems. To the contrary, for example, a particular classroom observation protocol or a particular system for tracking the placements and retention of program graduates might be utilized in a strong equity–centered evaluation system. However, this would depend on whether these evaluation tools were being used as part of a larger strong equity–centered evaluation approach that involved the authentic representation of minoritized families and community members, recognized the cultural values of minoritized groups, focused on redistribution of educational and other social and political resources, and worked at the level of structures and systems. When teacher preparation evaluation is strong equity–centered, as we call for in this

paper, the particular mechanisms, processes, and content used are jointly determined by relevant professional organizations, participants in teacher education programs and institutions, and members of local communities, schools, and families through a co-design process (Ishimaru et al., 2018, 2019). This means that evaluation tools cannot be completely pre-determined, but rather emerge from the "knowledge, priorities, and agendas" of students, families, and communities, (Ishimaru et al., 2019, p. 8), the goals of programs and participants, and appropriate notions of trustworthiness and validity. Along these lines, this work involves expanding "what counts" (Ishimaru, 2020; Ishimaru et al., 2018, 2019) in teacher preparation program evaluation by changing teacher preparation program evaluation metrics, policies, and practices to draw on the cultural values and shared knowledge and experience of students, families, and communities.

CONCLUSION

In light of the issues that we raise above, we propose that the NAEd project on *Evaluating and Improving Teacher Preparation Programs* call for a new approach and new models of teacher preparation evaluation with strong equity at their center, a proposal that we believe is consistent with the original prospectus for the NAEd project. To support this task, we recommend a set of principles and guidelines (see Box 1) organized according to the four dimensions of strong equity listed above and consistent with recent discussions of equity in the evaluation field (e.g., Center for Evaluation Innovation, 2017; Farrow & Morrison, 2019; Gates, 2020; Mertens & Zimmerman, 2015; NASEM, 2019; Schwandt & Gates, 2016, 2021; Wiggins & Sileo, 2020).

Ideally, this paper would conclude with examples of teacher preparation evaluation systems—at the federal, state, or professional levels—that are in keeping with the guiding principles outlined above. To our knowledge, however, evaluation *systems* of this kind do not exist in the United States, and although there are some evaluation systems in other developed countries consistent with some aspects of what we are calling for here, policy borrowing at the level of systems does not seem feasible. It should be noted, however, that although there are no preparation evaluation *systems* that make equity the centerpiece, there *are* many local programs designed to do so and are involved in efforts to assess their work. For example, to address the disconnect between teacher candidates and the communities they serve, which is related to the strong equity dimensions of representation and recognition, a number of preparation programs have endeavored to establish equitable relationships with community members most affected by inequities as co-teacher educators who are involved in decision-making about teacher preparation curricula, fieldwork experiences, teacher candidate evaluation, and program assessment.[3] Additionally, a number of programs have endeavored to establish and sustain

[3] See descriptions and analyses of community-based preparation programs (Murrell, 2000), such as the Schools Within the Context of Community teacher preparation program at Ball State University (Ball State University, 2017; Zygmunt & Clark, 2015); University of Washington's community-centered preparation programs wherein community-based educators share co-equal status as teacher educators (Guillen & Zeichner, 2018; Kretchmar & Zeichner, 2016; Zeichner et al., 2015); Loyola University Chicago's field-based teacher preparation program created in partnership with the Kateri urban Indigenous community organization (Lees, 2016); Lesley University's program that prepares educators to teach autistic students by positioning them as equals in the community (Keefe, 2015, 2016); and the University of California, Los Angeles, Center X's long-standing preparation program that evaluates teacher candidates on their activist skills in working with immigrant families and other minoritized students, families, and communities (Quartz, 2003).

BOX 1
Guiding Principles for a Strong Equity–Centered Approach to Teacher Preparation Evaluation

Dimension 1: Reframing evaluation

1. Establish *strong equity as an explicit goal and desired outcome* of teacher preparation evaluation, not a presumed by-product.
2. Build *attention to equity into the entire process of teacher preparation evaluation*, including establishing the purposes of the evaluation, deciding on how evidence will be generated and used and how validity will be defined, determining who will have an authentic voice in establishing the purposes and values that drive evaluation (including members of non-dominant communities served by the schools that partner with teacher preparation programs), and how the composition and diversity of the groups that are involved in evaluation policies and practices will be determined.
3. Utilize evaluation *tools and instruments* related to teacher preparation structures, program components, processes, curricula, and assessments that have the capacity to provide usable information *for ongoing programmatic self-examination and improvement with attention to issues of equity in all areas.*
4. Draw on the expertise and experience of an interdisciplinary task force that includes all relevant stakeholders in teacher preparation, including teacher education practitioners and members of non-dominant communities served by the schools that partner with preparation programs, to establish a set of *key teacher preparation equity indicators* to be used across teacher preparation evaluations.

Dimension 2: Redistribution of educational opportunities, access, and resources *and* resources beyond education

5. *Work at a structural/systems level* in teacher preparation evaluation that recognizes and addresses the multiple systemic and structural barriers—in addition to teacher quality—that produce and reproduce inequality in students' achievement and other school outcomes, such as poverty; inequities in school funding, school organization and support, family and community resources; institutionalized racism; and social policies and practices that maintain or exacerbate inequities related to health care, housing, transportation, jobs, law enforcement, and early childhood services.
6. Consider *problems and unequal outcomes and opportunities as the responsibility of "the system,"* not simply of individual actors, such as teacher educators, preparation programs, teachers, teacher candidates, and school-based teachers and leaders.

Dimension 3: Representation of multiple stakeholders

7. Ensure that all of those with a genuine stake in teacher preparation evaluation, including teacher educators, school-based educators, the families and community members served by the schools (including those from minoritized communities most affected by the inequities that exist), and members of professional organizations are authentically represented in purpose-setting, decision-making, evidence generation and interpretation, and determinations of consequences regarding teacher preparation evaluation.
8. Acknowledge and address power issues in evaluation and incorporate, as appropriate, approaches that are intentionally designed to share power and address power imbalances between external evaluators and those being evaluated, on one hand, and between teacher preparation programs and the communities they serve on the other hand.

continued

BOX 1 Continued

9. Focus on internal professional responsibility rather than imposed external accountability; external accountability agencies should be charged with supporting the capacity for strong internal accountability, supporting local innovation, and supporting the democratization of knowledge for teaching and teacher education.

Dimension 4: Recognition of cultural values that are not part of dominant institutionalized hierarchies

10. Recognize and draw on the perspectives, knowledge sources, and experiences of those most affected by the root causes of inequity (especially parents, families, and community members from minoritized communities).
11. Where appropriate, include evaluation models such as participatory evaluation, empowerment evaluation, Indigenous evaluation, and culturally responsive evaluation, all of which are intended to recognize and build on the cultural values of minoritized groups.

equitable partnerships between schools and universities (e.g., Burroughs et al., 2020; McDonald et al., 2014). Although these programs are clearly not teacher preparation evaluation systems, they are in keeping with the principles that we have proposed and can be informative.

Finally, the 2019 consensus study report of the National Academies of Sciences, Engineering, and Medicine (the National Academies), *Monitoring Educational Equity*, and particularly the notion of "key equity indicators" are also highly relevant to our recommendations here. Although the National Academies project focuses on K-12 education and not teacher preparation, it provides insights into how a group of scholars, researchers, and practitioners can reframe evaluation with equity at the center. The National Academies committee of experts in law, behavioral and social sciences, and measurement and statistics developed 16 key equity indicators focused on K-12 student outcomes as well as resources and opportunities, thus recommending a *system of indicators* with the capacity to examine disparities across racial, ethnic, linguistic, identified disability, and socioeconomic groups. As committee chair and legal scholar Christopher Edley, Jr., noted, "We think that the equity issue is so important and salient at this moment in time that a focus on educational equity deserves its own space, not simply as a piece of an existing set of data instruments." In concluding this paper, we echo Edley's sentiments. We believe that equity issues are so important in teacher preparation that it is essential to make strong equity central in new models for evaluating and improving teacher preparation. We recommend an inclusive approach to program evaluation, consistent with the National Academies' project noted above. This could include the development of key equity indicators for strong equity–centered teacher preparation evaluation, which could be informative for multiple teacher preparation constituencies, including teacher educators; their community, family, and school partners; professional organizations; state and federal policymakers; and other advocacy organizations.

REFERENCES

Alkin, M. (Ed.). (2004). *Evaluation roots: Tracing theorists views and influences* (2nd ed.). Sage Publications.

Alkin, M., & Christie, C. (2004). An evaluation theory tree. In Alkin, M. (Ed.), *Evaluation roots*: *Tracing theorists' views and influences* (2nd ed.). Sage Publications.

Alkin, M., & Christie, C. (2019). Theorists' models in action: A second look. *New Directions for Evaluation*, *2019* (163), 11-18.

Allen, M., Coble, C., & Crowe, E. (2014). *Building an evidence-based system for teacher preparation.* Teacher Preparation Analytics. http://caepnet.org/~/media/Files/caep/accreditation-resources/tpa-report-full.pdf?la=en.

Almy, S., Tooley, M., & Hall, D. (2013). *Preparing and advancing teachers and school leaders: A new approach for federal policy.* The Education Trust. https://edtrust.org/wp-content/uploads/2013/10/Preparing_and_Advancing_0.pdf.

Ambrosio, J. (2013). Changing the subject: Neoliberalism and accountability in public education. *Educational Studies*, *49*(4), 316-333.

American Association of Colleges for Teacher Education. (2020). *edTPA participation map*. http://edtpa.aacte.org/state-policy.

American Educational Research Association. (2015). AERA statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher*, *44*(8), 448-452.

Anderson, L. (2019). Private interests in a public profession: Teacher education and racial capitalism. *Teachers College Record*, *121*(4), 1-18.

Andrews, D. J. C., Brown, T., Castillo, B. M., Jackson, D., & Vellanki, V. (2019a). Beyond damage-centered teacher education: Humanizing pedagogy for teacher education: Humanizing pedagogy for teacher educators and preservice teachers. *Teachers College Record*, *121*(6), 1-28.

Andrews, K., Parekh, J., & Peckoo, S. (2019b). *How to embed a racial and ethnic equity perspective in research: Practical guidance for the research process* [Working paper]. Child Trends. https://www.childtrends.org/wp-content/uploads/2019/09/RacialEthnicEquityPerspective_ChildTrends_October2019.pdf.

Association for Advancing Quality in Educator Preparation. (2020). *AAQEP collaboration & innovation in quality assurance.* Webinar, June 1, 2020. Association for Advancing Quality in Educator Preparation.

Association for Advancing Quality in Educator Preparation. (2021). *Guide to AAQEP accreditation*. https://aaqep.org/files/2020-Guide-to-AAQEP-Accreditation.pdf.

Au, W. (2016). Meritocracy 2.0: High-stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educational Policy*, *30*(1), 39-62.

Bales, B. (2006). Teacher education policies in the United States: The accountability shift since 1980. *Teaching and Teacher Education*, *22*(4), 395-407.

Ball State University. (2017). *Schools within the context of community*. http://cms.bsu.edu/academics/collegesanddepartments/teachers/departments/elementaryed/academics/immersive-learning/scc.

Bell, C. A., & Youngs, P. (2011). Substance and show: Understanding responses to teacher education programme accreditation processes. *Teaching and Teacher Education*, *27*(2), 298-307.

Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value added models.* Educational Testing Service.

Brown, K. D. (2013). Trouble on my mind: Toward a framework of humanizing critical sociocultural knowledge for teaching and teacher education. *Race Ethnicity and Education*, *16*(3), 316-338.

Burroughs, G., Lewis, A., Battey, D., Curran, M., Hyland, N. E., & Ryan, S. (2020). From mediated fieldwork to co-constructed partnerships: A framework for guiding and reflecting on P-12 school-university partnerships. *Journal of Teacher Education*, *71*(1), 122-134.

Carinci, J. E., Meyer, S. J., & Jackson, C. (Eds.). (2020). *Linking teacher preparation program design and implementation to outcomes for teachers and students.* Information Age Publishing.

Carter, P. (2013). Student and school cultures and the opportunity gap: Paying attention to academic engagement and achievement. In P. Carter & K. Welner (Eds.), *Closing the opportunity gap: What America must do to give every child an even chance* (pp. 143-155). Oxford University Press.

Christie, C. A., & Alkin, M. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation*, *34*(3), 131-135.

Christie, C., & Alkin, M. (2013). An evaluation theory tree. In M. Alkin (Ed.), *Evaluation roots*: *A wider perspective of theorists' views and influences* (2nd ed.). Sage Publications.

Cochran-Smith, M. (1995a). Color blindness and basket making are not the answers. *American Educational Research Journal*, *32*(3), 493-522.

Cochran-Smith, M. (1995b). Uncertain allies: Understanding the boundaries of race and teaching. *Harvard Educational Review*, *63*(4), 541-570.

Cochran-Smith, M. (2005). The new teacher education: For better or for worse? *Educational Researcher*, *34*(7), 3-17.

Cochran-Smith, M. (2010). Toward a theory of teacher education for social justice. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second international handbook of educational change* (pp. 445-467). Springer Science & Business Media.

Cochran-Smith, M., Carney, M. C., Keefe, E. S., Burton, S., Chang, W-C., Fernández, M. B., Miller, A., Sánchez, J. G., & Baker, M. (2018). *Reclaiming accountability in teacher education.* Teachers College Press.

Cochran-Smith, M., & Fries, K. (2001). Sticks, stones, and ideology: The discourse of reform in teacher education. *Educational Research*, *30*(8), 3-15.

Cochran-Smith, M., & Keefe, E. S. (in press). Strong equity: Repositioning equity in teacher education. *Teachers College Record.*

Cochran-Smith, M., Piazza, P., & Power, C. (2013). The politics of accountability: Assessing teacher education in the United States. *The Educational Forum*, *77*(1), 6-27.

Cochran-Smith, M., Stern, R., Sánchez, J. G., Miller, A., Keefe, E. S., Fernández, M. B., Chang, W. C., Carney, M. C., Burton, S., & Baker, M. (2016). *Holding teacher preparation accountable: A review of claims and evidence.* National Education Policy Center. https://nepc.info/sites/default/files/pb_cochran-smith_teacher_prep_0.pdf.

Coggshall, J. D., Bivona, L., & Reschly, D. J. (2012). *Evaluating the effectiveness of teacher preparation programs for support and accountability.* National Comprehensive Center for Teacher Quality. https://files.eric.ed.gov/fulltext/ED543773.pdf.

Cohen-Vogel, L., & Hunt, H. (2007). Governing quality in teacher education: Deconstructing federal text and talk. *American Journal of Education*, *114*(1), 137-163.

Colvin, R. L. (2005). A new generation of philanthropists. In F. Hess (Ed.), *With the best of intentions: How philanthropy is reshaping K-12 education* (pp. 21-48). Harvard Education Press.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Houghton Mifflin.

Council of Chief State School Officers. (2012). *Our responsibility, our promise: Transforming educator preparation and entry into the profession*. Task Force on Educator Preparation and Entry into the Profession. https://www.ccsso.org/sites/default/files/2017-10/Our%20Responsibility%20Our%20Promise_2012.pdf.

Council of Chief State School Officers. (2017). *Transforming educator preparation: Lessons learned from leading states*. https://ccsso.org/sites/default/files/2017-11/CCSSO%20Educator%20Preparation%20Playbook.pdf.

Council of Chief State School Officers. (2018). *Measuring what matters: Recommendations from states in the network for transforming educator preparation*. http://www.ccsso.org/sites/default/files/2018-02/Measuring%20What%20Matters.pdf.

Crowe, E. (2010). *Measuring what matters: A stronger accountability model for teacher education*. Center for American Progress. https://files.eric.ed.gov/fulltext/ED518518.pdf.

Crowe, E. (2011a). *Race to the Top and teacher preparation*. Center for American Progress.

Crowe, E. (2011b). *Getting better at teacher preparation and state accountability strategies, innovations, and challenges under the federal Race to the Top program.* Center for American Progress.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1).

Darling-Hammond, L., Newton, S. P., & Wei, R. C. (2012). *Developing and assessing beginning teaching effectiveness: The potential of performance assessments*. Stanford Center for Opportunity Policy in Education.

Datta, L. (2011). Politics and evaluation: More than methodology. *American Journal of Education*, *32*(2), 273-294.

Deans for Impact. (2016). *From chaos to coherence: A policy agenda for accessing and using outcomes data in educator preparation*. http://www.deansforimpact.org/wp-content/uploads/2016/11/From_Chaos_to_Coherence.pdf.

Duncan, A. (2009). *Teacher preparation: Reforming the uncertain profession*. U.S. Department of Education. http://www2.ed.gov/news/pressreleases/2009/10/10222009a.html.

Easton, J. Q. (2008, October). *Goals and aims of value-added modeling: A Chicago perspective.* University of Chicago Consortium on School Research.

Economic Policy Institute. (2010). *Problems with the use of student test scores to evaluate teachers.* https://www.epi.org/publication/bp278.

Education Deans for Justice and Equity. (2019a). *Seven trends in U.S. teacher education, and the need to address systemic injustices*. National Education Policy Center. https://nepc.colorado.edu/sites/default/files/publications/FYI%20Ed%20Deans_1.pdf.

Education Deans for Justice and Equity. (2019b). *A framework for assessment and transformation* ("The EDJE Framework" v.6.20.2019). https://drive.google.com/file/d/1AVVkmzPxEMAt_g-BPYyNVHgseylmREnN/view.

The Education Trust. (2017). *Tackling gaps in access to strong teachers: What state leaders can do.* https://edtrust.org/resource/tackling-gaps-access-strong-teachers.

Equitable Evaluation Initiative. (2017). *Equitable evaluation framing paper*. www.equitableeval.org.

Farrow, F., & Morrison, S. (2019). *Placing equity concerns at the center of knowledge development.* Center for the Study of Social Policy.

Feuer, M. J., Floden, R. E., Chudowsky, N., & Ahn, J. (2013). *Evaluation of teacher preparation programs: Purposes, methods, and policy options.* National Academy of Education.

Fraser, N. (2003). Social justice in an age of identity politics: Redistribution, recognition and participation. In N. Fraser & A. Honneth (Eds.), *Redistribution or recognition: A political-philosophical debate* (pp. 7-109). Verso.

Fraser, N. (2009). *Scales of justice. Reimagining political space in a globalizing world*. Columbia University Press.

Fullan, M., Rincon-Gallardo, S., & Hargreaves, A. (2015). Professional capital as accountability. *Education Policy Analysis Archives*, *23*(15), 1-22.

Fuller, E. J. (2015). Shaky methods, shaky motives: A critique of the National Council of Teacher Quality's review of teacher preparation programs. *Journal of Teacher Education*, *65*(1), 63-77.

Gates, E. F. (2017). Toward valuing with critical systems heuristics. *American Journal of Evaluation*, *39*(2), 201-220.

Gates, E. (2020). *Toward developing an action-guiding theory for equity in educational evaluation.* Unpublished paper. Boston College.

Goldhaber, D. (2013). What do value-added measures of teacher preparation programs tell us? Carnegie Foundation for the Advancement of Teaching.

Goldhaber, D. (2019). Evidence-based teacher preparation: Policy context and what we know. *Journal of Teacher Education*, *70*(2), 90-101.

Goldhaber, D., Naito, N., & Theobald, R. (2019). Making the most of student teaching: The importance of mentors and scope for change. *Education Finance and Policy*, *15*(3), 581-591.

Gollnick, D. (1992). Multicultural education: Policies and practices in teacher education. In C. Grant (Ed.), *Research and multicultural education: From the margins to the mainstream* (pp. 218-239). Falmer Press.

Grant, C. (2008). Teacher capacity: Introduction to the section. In M. Cochran-Smith, S. Fiemen-Nemser, J. McIntype, & K. Demers (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (3rd ed., pp. 127-133). Routledge.

Greene, J. C. (2006). Evaluation, democracy, and social change. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *Handbook of evaluation: Policies, programs, and practices* (pp. 118-140). Sage Publications. https://learningpolicyinstitute.org/product/teacher-residency.

Guillen, L., & Zeichner, K. (2018). A university-community partnership in teacher education from the perspectives of community-based teacher educators. *Journal of Teacher Education*, *69*(2), 140-153.

Haertel, E. H. (2013). *Reliability and validity of inferences about teachers based on student test scores*. Educational Testing Service.

Hall, M. (2018). Evaluation's race problem in the United States: A Call to Action for the profession and the *American Journal of Evaluation. American Journal of Evaluation*, *39*(4), 569-583.

Haugen, J. S., & Chouinard, J. A. (2019). Transparent, translucent, opaque: Exploring the dimension of power in culturally responsive evaluation contexts. *American Journal of Evaluation*, *40*(3), 376-394.

Hess, F. (2005). Introduction. In F. Hess (Ed.), *With the best of intentions: How philanthropy is reshaping K-12 education* (pp. 1-17). Harvard Education Press.

Hood, S., Hopson, R., & Kirkhart, K. (2015). Culturally responsive evaluation: Theory, practice and future implications. In K. Newcomer & H. Hatry (Eds.), *Handbook of practical program evaluation* (pp. 281-317). Jossey-Bass.

Horn, I. S., & Kane, B. D. (2019). What we mean when we talk about teaching: The limits of professional language and possibilities for professionalizing discourse in teachers' conversations. *Teachers College Record*, *121*(4), 1-30.

House, E. R., & Howe, K. R. (2000). Deliberative democratic evaluation in practice. In D. Stufflebeam, T. Kellaghan, & G. Maddaus (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 409-421). Springer.

Iasevoli, B. (2017, March 28). Trump signs bill scrapping teacher-prep rules [Blog post]. http://blogs.edweek.org/edweek/teacherbeat/2017/03/trump_signs_bill_scrapping_tea.html.

Ishimaru, A. M. (2020). *Just schools: Building equitable collaborations with families and communities.* Teachers College Press.

Ishimaru, A. M., Bang, M., Valladares, M. R., Nolan, C. M., Tavares, H., Rajendran, A., & Chang, K. (2019). *Recasting families and communities as co-designers in education in tumultuous times*. National Education Policy Center. https://nepc.colorado.edu/publication/family-leadership.

Ishimaru, A. M., Rajendran, A., Nolan, C. M., & Bang, M. (2018). Community design circles: Co-designing justice and well-being in family-community-research partnerships. *Journal of Family Diversity in Education*, *3*(2), 38-63.

Johnson, S. M., Birkeland, S. E., & Peske, H. G. (2005). Life in the fast track: How states seek to balance incentives and quality in alternative teacher certification programs. *Educational Policy*, *19*(1), 63-89.

Kantor, H., & Lowe, R. (2013). Educationalizing the welfare state and privatizing education the evolution of social policy since the New Deal. In P. L. Carter & K. G. Welner (Eds.), *Closing the opportunity gap: What America must do to give every child an even chance* (pp. 25-39). Oxford University Press.

Keefe, E. S. (2015, June). *Autism Spectrum Disorder: Through a democratic lens* [Presentation]. Wheelock College.

King, J. (2008). Critical and qualitative research in teacher education: A blues epistemology, a reason for knowing for cultural well-being. In M. Cochran-Smith, S. Feiman Nemser, & J. McIntyre (Eds.), *Handbook of research on teacher education: Enduring issues in changing contexts* (pp. 1094-1135). Lawrence Erlbaum Associates, Inc.

Kornfeld, J., Grady, K., Marker, P. M., & Ruddell, M. R. (2007). Caught in the current: A self-study of state-mandated compliance in a teacher education program. *Teachers College Record*, *109*(8), 1902-1930.

Kretchmar, K., & Zeichner, K. (2016). Teacher prep 3.0: A vision for teacher education to impact social transformation. *Journal of Education for Teaching*, *42*(4), 417-433.

Kumashiro, K. (2015). *Review of proposed 2015 federal teacher preparation regulations.* National Education Policy Center. http://nepc.colorado.edu/thinktank/review-proposed-teacher-preparation.

Ladson-Billings, G. (1999). Preparing teachers for diverse student populations: A critical race theory perspective. In A. Iran-Nejad & D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 211-248). American Educational Research Association.

Lather, P. (1992). Critical frames in educational research: Feminist and post-structural perspectives. *Theory Into Practice*, *31*(2), 87-99.

Lees, A. (2016). Roles of urban Indigenous community members in collaborative field-based teacher preparation. *Journal of Teacher Education*, *67*(5), 363-378.

Lincove, J. A., Osborne, C., Mills, N., & Bellows, L. (2015). Teacher preparation for profit or prestige: Analysis of a diverse market for teacher preparation. *Journal of Teacher Education*, *66*(5), 415-434.

Lipman, P. (2011). Neoliberal education restructuring dangers and opportunities of the present crisis. *Monthly Review*, *63*(3), 114-127.

Mathis, W., & Trujillo, T. M. (2016). *Learning from the federal market based reforms: Lessons for ESSA.* National Education Policy Center.

McDonald, M., Kazemi, E., Kelley-Peterson, E., Mikolasy, K., Thompson, J., Valencia, S., & Windschitl, M. (2014). Practice makes practice: Learning to teach in teacher education. *Peabody Journal of Education*, *89*(4), 500-515.

McKinsey & Company. (2007, September). *How the world's best performing school systems come out on top.* http://mckinseyonsociety.com/downloads/reports/Education/Worlds_School_Systems_Final.pdf.

Mehta, J. (2013). How paradigms create politics: The transformation of American educational policy, 1980-2001. *American Educational Research Journal*, *50*(2), 285-324.

Mertens, D., & Wilson, A. (2012). *Program evaluation theory and practice: A comprehensive guide.* The Guilford Press.

Mertens, D., & Wilson, A. (2019). *Program evaluation theory and practice: A comprehensive guide* (2nd ed.). The Guilford Press.

Mertens, D., & Zimmerman, H. (2015). A transformative framework for culturally responsive evaluation. In S. Hood, R. Hopson, & H. Freirson (Eds.), *Continuing the journey to reposition culture and cultural contexts in evaluation theory and practice* (pp. 275-287). Information Age Publishing.

The MET Project. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains* [Research paper]. Bill & Melinda Gates Foundation.

Milner, H. R., IV., Pearman, F. A., III, & McGee, E. O. (2013). Critical race theory: Interest convergence, and teacher education. In M. Lynn & A. D. Dixson (Eds.), *Handbook of critical race theory in education* (pp. 339-354). Routledge.

Murrell, P. (2000). Community teachers: A conceptual framework for preparing exemplary urban teachers. *The Journal of Negro Education*, *69*(4), 338-348.

National Academies of Sciences, Engineering, and Medicine. (2019). *Monitoring Educational Equity*. The National Academies Press.

National Council for Accreditation of Teacher Education Blue Ribbon Commission. (2010). *Transforming teacher education through clinical practice: A national strategy to prepare effective teachers.* Report of the Blue Ribbon Panel on Clinical Preparation and Partnerships for Improved Student Learning. National Council for Accreditation of Teacher Education Blue Ribbon Commission.

National Research Council. (2010). *Preparing teachers: Building evidence for sound policy.* The National Academies Press.

Nieto, S. (2010). *Language, culture, and teaching: Critical perspectives.* Routledge.

Noell, G. H., Burns, J. M., & Gansle, K. A. (2019). Linking student achievement to teacher preparation: Emergent challenges in implementing value added assessment. *Journal of Teacher Education*, *70*(2), 128-138.

O'Neil, O. (2013). Intelligent accountability in education. *Oxford Review of Education*, *39*(1), 4-16.

Organisation for Economic Co-operation and Development. (2005). *Teachers matter: Attracting, developing and retaining effective teachers*. Organisation for Economic Co-operation and Development.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage.

Philip, T. (2019). Principled improvisation to support novice teacher learning. *Teachers College Record*, *121*(4), 1-28.

Philip, T. M., Souto-Manning, M., Anderson, L., Horn, I., Carter Andrew, D. J., Stillman, J., & Varghese, M. (2018). Making justice peripheral by constructing practice as "core": How the increasing prominence of core practices challenges teacher education. *Journal of Teacher Education*, *70*(3), 251-264.

Phillips, D. C., & Burbules, N. C. (2001). *Postpositivism and educational research.* Rown & Littlefield.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109-119.

Plecki, M., Elfers, A. M., & Nakamara, Y. (2012). Using evidence for teacher education program improvement and accountability: An illustrative case of the role of value-added measures. *Journal of Teacher Education*, *63*(5), 318-334.

Quartz, K. H. (2003). Too angry to leave: Supporting new teachers' commitment to transform urban schools. *Journal of Teacher Education, 54*(2), 99-111.

Ronfeldt, M. (2021). *Links among teacher preparation, retention, and teaching effectiveness*. National Academy of Education Committee on Evaluating and Improving Teacher Preparation Programs. National Academy of Education.

Ronfeldt, M., Brockman, S. L., & Campbell, S. L. (2018). Does cooperating teachers' instruction effectiveness improve preservice teachers' future performance? *Educational Researcher*, *47*(7), 405-418.

Schwandt, T. A., & Gates, E. F. *(2016). What can evaluation do? An agenda for evaluation in service of an equitable society. In* S. Donaldson & R. Picciotto (Eds.), *Evaluation for an equitable society* (pp. 67-81). Information Age Publishing.

Schwandt, T. A., & Gates, E. F. (2021). *Evaluating and valuing in social research*. Guilford Press.

Scott, J. (2016). The politics of market-based education reform. In M. Matthis & T. Trujullio (Eds.), *Learning from the federal market-based reforms* (pp. 9-37). Information Age Publishing.

Sleeter, C. (2001). Preparing teachers for culturally diverse schools: Research and the overwhelming presence of whiteness. *Journal of Teacher Education*, *52*(2), 94-106.

Sleeter, C. (2009). Teacher education, neoliberalism, and social justice. In W. Ayres, T. Quinn, & D. Stovall (Eds.), *The handbook of social justice in education* (pp. 629-642). Taylor and Francis.

Sleeter, C. (2017). Critical race theory and the whiteness of teacher education. *Urban Education*, *52*(2), 155-169.

Souto-Manning, M. (2019). Transforming university-based teacher education: Preparing asset-, equity-, and justice-oriented teachers within the contemporary political context. *Teachers College Record*, *121*(4), 1-26.

Stillman, J., & Beltramo, J. L. (2019). Exploring Freirean culture circles and Boalian theatre as pedagogies for preparing asset-oriented teacher educators. *Teachers College Record*, *121*(4), 1-32.

Taubman, P. M. (2009). *Teaching by numbers: Deconstructing the discourse of standards and accountability in education*. Routledge.

Thomas, V. G., & Campbell, P. B. (2020). *Evaluation in today's world: Respecting diversity, improving quality, and promoting usability.* Sage Publishing.

TPI-US (2020, December 7). *TPI-US Better prepared teachers means better outcomes for students*. https://www.tpius.org.

U.S. Department of Education. (2002). *Meeting the highly qualified teachers challenge: The Secretary's annual report on teacher quality.* https://files.eric.ed.gov/fulltext/ED513876.pdf.

U.S. Department of Education. (2003). *Meeting the highly qualified teachers challenge: The Secretary's second annual report on teacher.* https://files.eric.ed.gov/fulltext/ED498864.pdf.

U.S. Department of Education. (2011). *Our future, our teachers: The Obama administration's plan for teacher education reform and improvement*. ed.gov/sites/default/files/our-future-our-teachers.pdf.

U.S. Department of Education. (2014). *Teacher preparation issues; Proposed rule, 2014, 79, 232 Fed. Reg. 71820 (December 3, 2014)*. https://www.govinfo.gov/content/pkg/FR-2014-12-03/html/2014-28218.htm.

Valli, L., & Rennert-Ariev, P. (2002). New standards and assessments? Curriculum transformation in teacher education. *Journal of Curriculum Studies*, *34*(2), 201-225.

Villegas, A., & Lucas, T. (2004). Diversifying the teacher workforce: A retrospective and prospective account. In M. Smylie & D. Miretzky (Eds.), *Developing the teacher workforce: The 103rd yearbook of the National Society for the Study of Educatio*n (pp. 70-104). University of Chicago Press.

Von Hippel, P. T., & Bellows, L. (2018). Rating teacher-preparation programs: Can value added make useful distinctions? *Education Next*, *18*(3), 35-41.

Walby, S. (2007). Complexity theory, systems theory, and multiple intersecting social inequalities. *Philosophy of the Social Sciences*, *37*(4), 449-470.

Weiss, C. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, *19*(1), 21-33.

Wiggins, M. E., & Sileo, A. (2020). *What's the role of equity in evaluation policy?* The Forum for Youth Investment.

World Bank. (2010, May 4). *Teacher politics around the world: Objectives, rationale, methodological approach and products.* http://siteresources.worldbank.org/EDUCATION/Resources/Vegasetal__Teacherpolicies aroundtheworld.draft.pdf.

Worrell, F., Brabeck, M., Dwyer, C., Geisinger, K., Marx, R., Noell, G., & Pianta, R. (2014). *Assessing and evaluating teacher preparation programs*. American Psychological Association. https://www.apa.org/ed/schools/teaching-learning/teacher-preparation-programs.pdf.

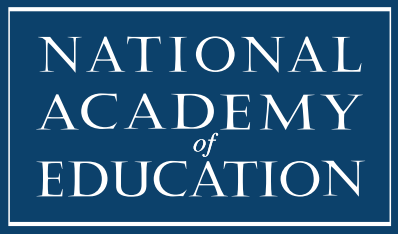
Zeichner, K. (2003). The adequacies and inadequacies of three current strategies to recruit, prepare, and retain the best teachers for all students. *Teachers College Record*, *105*(3), 490-515.

Zeichner, K. (2009). *Teacher education and the struggle for social justice.* Routledge.
Zeichner, K. (2018). *The struggle for the soul of teacher education*. Routledge.
Zeichner, K., & Hoeft, K. (1996). Teacher socialization for cultural diversity. In J. Sikula, T. J. Buttery, & E. Guyton (Eds.), *Handbook of research on teacher education* (pp. 525-547). MacMillan.
Zeichner, K., Payne, K. A., & Brayko, K. (2015). Democratizing teacher education. *Journal of Teacher Education*, *66*(2), 122-135.
Zeichner, K., & Peña-Sandoval, C. (2015). Venture philanthropy in teacher education policy in the U.S.: The role of the New Schools Venture Fund. *Teachers College Record*, *117*(6), 1-24.
Zygmunt, E., & Clark, P. (2015). *Transforming teacher education for social justice.* Teachers College Press.

AUTHOR BIOGRAPHIES

Marilyn Cochran-Smith is the Cawthorne Professor of Teacher Education at the Lynch School of Education and Human Development at Boston College. She has been a teacher education practitioner, scholar, and researcher for more than 40 years. Her research interests include practitioner inquiry and teacher education research, practice, and policy with a focus on social justice and equity. Cochran-Smith is a frequent speaker nationally and internationally with recent keynotes in Australia, England, Ireland, Israel, Malta, New Zealand, and Norway. She has written 10 books, 7 of which have won national awards, and more than 200 articles, chapters, and editorials. Her co-authored book, *Reclaiming Accountability in Teacher Education* (Teachers College Press, 2018), won the 2020 Best Book Award from the American Association of Colleges of Teacher Education and the American Educational Research Association's (AERA's) Division K 2019 Distinguished Contributions to Research Award. Cochran-Smith was also a 2020 winner of the Spencer Foundation's Mentorship Award as well as the 2018 winner of the AERA Division K Lifetime Achievement Award. She is currently the principal investigator for a Spencer Foundation–funded study of U.S. "new graduate schools of education," which prepare teachers and award master's degrees, but are not university based or affiliated. Cochran-Smith is an elected member of the National Academy of Education, a former president of AERA, an AERA Fellow, and an elected member of the Laureate Chapter of the Kappa Delta Pi National Education Honorary Society.

Emilie Mitescu Reagan is an associate professor of education in the Claremont Graduate University (CGU) School of Educational Studies. Reagan's research focuses on social justice–oriented teacher education policy and practice, using primarily quantitative and mixed-methods research. Specifically, she examines how teacher education programs and policies support preservice teacher learning and practice. Additionally, she critically analyzes assessment and accountability systems in teacher education. As part of her research, Reagan has secured external funding to develop programs and conduct research on multiple processes and outcomes of teacher education, including a multi-million dollar U.S. Department of Education Teacher Quality Partnership grant to design and implement a rural teacher residency program (as co-principal investigator [PI]) and a grant from the Spencer Foundation to conduct a multi-institutional, mixed-methods study examining the relationship between teacher candidate performance assessments and novice teacher learning (as PI). She has published articles in peer-reviewed journals including *Teaching and Teacher Education*, *Action in Teacher Education*, *Urban Education*, and *Education Policy Analysis Archives*. Reagan is also the president of the New England Educational Research Organization (2020-2022). Prior to joining the CGU faculty, Reagan was an associate professor in the University of New Hampshire Department of Education. Reagan began her career in education as a fifth-grade teacher in Philadelphia, Pennsylvania.



The **National Academy of Education (NAEd)** advances high-quality research to improve education policy and practice. Founded in 1965, the NAEd consists of U.S. members and international associates who are elected on the basis of scholarship related to education. The Academy undertakes research studies to address pressing educational issues and administers professional development fellowship programs to enhance the preparation of the next generation of education scholars.

naeducation.org