Using a Participant Pool to Gather Data in a Teacher Education Program: The Course of One School's Efforts

Peter D. Wiens
University of Wisconsin-Parkside

Introduction

A decade ago, the Carnegie Corporation of New York, along with the Ford and Annenberg Foundations, undertook a reform initiative, *Teachers for a New Era* (TNE). The goal of the TNE was to "stimulate construction of excellent teacher education programs at selected colleges and universities" (Carnegie Corporation of New York, 2001, p. 1). As part of this project, one mid-Atlantic university designed and implemented a participant data pool (PDP) to collect and manage data on teacher education students. The purpose of the PDP was to use the data to better understand whether the teacher education program was, indeed, an excellent program and to further stimulate knowledge creation about teacher education and preservice teachers.

Ten years later, it is now time to examine the implementation of a PDP. In reflecting on the creation of PDP, McNergney and Imig (2006) wondered whether it was feasible to conduct teacher education program evaluations within and across settings. To this end, this study sought to examine the implementation of the PDP over the past ten years and to address McNergney and Imig's question.

Peter D. Wiens is an assistant professor with the Institute of Professional Educator Development at the University of Wisconsin-Parkside, Kenosha, Wisconsin. His e-mail address is wiens@uwp.edu

Review of the Literature

Research has consistently shown that teachers are the most important in-school factor in determining student achievement (Nye, Konstantopoulos, & Hedges, 2004; Sanders & Horn, 1998). In the last decade, teacher education programs have come under increased pressure to demonstrate that they are providing an important contribution to the development of effective teachers. The federal government has put a focus on teacher performance and teacher education through *Race to the Top*, which requires evidence of successful teacher education (U.S. Department of Education, 2009). In a broad condemnation of traditional teacher education, U.S. Secretary of Education Duncan stated that most university and college teacher education programs do a mediocre job of preparing teachers (Medina, 2009). Further, a report by the National Council on Teacher Quality (NCTQ) has deemed traditional teacher education "an industry of mediocrity" (Greenberg, McKee, & Walsh, 2013, p. 1). The report has received a great deal of attention because the NCTQ claims to have done what teacher education research has failed to do—empirically assess teacher education programs and compare across contexts.

High-profile alternative teacher preparation programs, such as Teach For America, have also put pressure on the traditional teacher education establishment to demonstrate the importance of its role in the educational system. However, empirical research has not provided systematic evidence of the efficacy of traditional teacher preparation. Instead, studies that have used statistical methods to control for differences in teacher placements have shown either small or no differences in the effectiveness of alternatively certified and traditionally certified teachers (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2005; Kane, Rockoff, & Staiger, 2006). In the face of such criticisms and evidence, teacher education programs need to make the argument that they provide an important role in training effective teachers.

Challenges of Studying Teacher Education

Research on teacher education has lacked a common agenda or methodology across situations and researchers (Grossman & McDonald, 2008). This has made it difficult to create a general knowledge of effective teacher education. An important challenge to studying teacher education's impact on student learning is the difficulty of connecting student outcomes to teacher education programs. There are many variables that play a role in student learning, and isolating the teacher characteristics or behaviors that make a difference in student learning is difficult (Goldhaber, 2008). Identifying which teacher characteristics are a result of a teacher educa-

tion program adds complexity and makes it challenging for researchers to estimate the effects of teacher preparation.

There is also great variety in teacher education programs (Wilson, Floden, & Ferrini-Mundy, 2002). National accreditation agencies, such as the National Council for the Accreditation of Teacher Education (NCATE), and state departments of education set guidelines for teacher education programs. Within these guidelines, however, individual teacher preparation programs have substantial latitude in determining how to educate future teachers. As a result, the number of classes required and the content of those classes vary significantly from program to program (Constantine et al., 2009).

Floden (2008) noted that much of the teacher education faculty at research universities take on heavy teaching loads, which leaves little time for these faculty members to engage in rigorous, large-scale research projects, including research on teacher education students and their growth. Alongside the lack of time for research among teacher education faculty is a general lack of money from grant organizations (Cochran-Smith & Zeichner, 2005). This combination has led teacher education research to lag behind other fields of educational inquiry.

Analyzing Growth in Teacher Education

One method for measuring effective teaching is the use of value-added models (VAMs) that use student achievement growth scores to estimate teacher effectiveness. VAMs can be used to compute teacher effects and then connect these effects to the teacher education programs from which the teacher graduated to derive an effectiveness estimate of the teacher education program. Several states, including Louisiana, Florida, and North Carolina, have begun to use VAMs in this manner. Although researchers have been using VAMs to estimate in-service teacher quality, it is only recently that researchers also have started to use VAMs to judge the effectiveness of teacher preparation (e.g. Boyd, Grossman, Lankford, Loeb, & Wykoff, 2009).

VAMs have received quite a lot of attention but remain controversial because they cannot show the value added of a teacher preparation program (Baker et al., 2010; Glazerman et al., 2010). They can show only the effectiveness of graduates of specific teacher education programs. VAMs cannot be used to measure preservice teachers at the beginning of teacher education programs, to understand the growth in teacher competencies, because the preservice teachers are not teaching in the field when they begin their programs. Therefore, VAMs cannot be used to distinguish between effective teacher education programs and teacher education programs that are effective at recruiting highly effective teachers.

While using value-added statistical models to assess teacher effectiveness is relatively new, for years, teacher education programs have used formal and informal methods to understand the impact of programs on students. Such methods are also an important part of the accreditation process, whereby the use of multiple measures of student growth are encouraged (NCATE, 2008).

Typically, assessments of student learning are collected and analyzed by the teacher education program itself. The specific measures and the constructs that they assess are decided upon by each individual teacher education program. NCATE endorses multiple measures of student growth and performance (NCATE, 2011). These measures include lesson plans, evaluations of content knowledge, student teaching, and portfolios, among others. Taken together, these measures may demonstrate a form of valued added for the program in terms of teacher education student growth and provide evidence to guide program revisions.

Conceptualizing a Data Collection System

The gold standard for determining the effectiveness of a teacher education program is to make the connection between the program and the learning of P-12 students in the classrooms of graduates of the program (Wiens, 2012b). Boyd, Goldhaber, Lankford, and Wyckoff (2007) noted, however, that this has not yet been accomplished. McNergney and Imig (2006) recommend, instead, that teacher education programs may benefit from focusing on short-term measures that can be administered several times and analyzed within a teacher education setting.

Because teachers must possess a variety of knowledge and skills to be successful (Brophy, 1999; Darling-Hammond, 2006a), any data collection system needs to include multiple measures to capture preservice teacher learning and growth (Cochran-Smith & Ziechner, 2005). A comprehensive data collection system would evaluate preservice teacher learning through multiple lenses. Darling-Hammond (2006b) stated, "Teacher educators are seeking to develop strategies for assessing the results of their efforts—strategies that appreciate the complexity of teaching and learning and that provide a variety of lenses on the process of learning to teach" (p. 120).

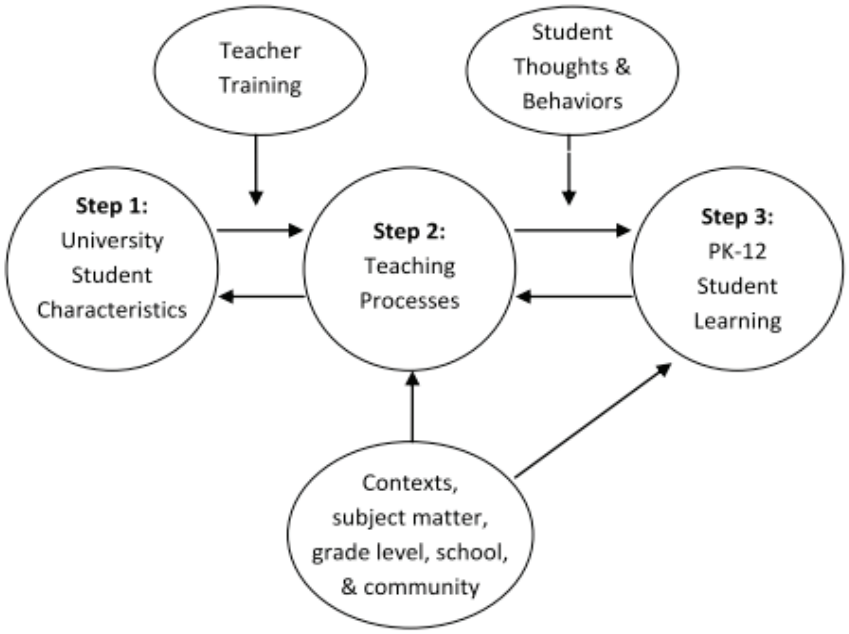
The data management and collection described in this paper is founded on the process-product conceptual model (Mitzel, 1960). McNergney and Imig (2006) developed a model of how this university constructed the teacher education research and evaluation formula (Figure 1). In Step 1, teacher education students enter the program with a certain set of characteristics. These characteristics interact with the teacher training provided by the university that leads to teaching processes.

Step 2 includes the practices of teaching. These are the curriculum and instruction that a teacher implements in his or her classroom. Step 3, PK-12 student learning products, is a result of teaching processes, student thoughts and behaviors, and contexts. The bidirectional arrows indicate the iterative nature of teacher education.

McNergney and Imig (2006) presented three groups of guiding questions to be used in the creation of a PDP:

- What are the intellectual, physical, emotional, and social capacities of pre-service teachers that are likely to influence their teaching behavior? How well have they mastered the content they will teach? Do pre-service teachers have the proclivity to continue learning?
- How do teachers in preparation demonstrate teaching behavior that has been shown to relate to or cause PK-12 student learning? Can they plan instruction, implement these plans, and assess their work in ways that are likely to influence students' performances?
- How do pre-service teachers' students think about teaching and

Figure 1
A Model for Teacher Education Research and Evaluation
(McNergney & Imig, 2008)



> learning—are they motivated, challenged, and reinforced by teachers' actions? What do pre-service teachers' students learn? (p. 4)

A comprehensive model for analyzing teacher education value-added would answer all of these questions.

The Participant Data Pool

In an effort to answer McNergney and Imig's (2006) questions and to understand student achievement and demonstrate growth under accountability requirements, the university has implemented a PDP. The PDP was designed as a means to collect, manage, and analyze data on the school's teacher education students. The PDP is broadly defined at the university as the data collected from survey and performance measures as well as the availability of participants for other research projects. All teacher education students are required to contribute to the PDP by participating in a set number of research activities each year in which they are enrolled in the program. Alternative assignments to participating in research studies are also offered; however, no student has elected to complete an alternative assignment. Students are held accountable for participation through classes in which they take part in the teacher education program.

The PDP, which constitutes only one part of the program's evaluation process, includes surveys and performances measures that have been used widely by researchers in the education literature. Table 1 presents the measures employed in the PDP. Survey measures were chosen by university faculty and researchers based on their widely accepted use, while performance measures were chosen by teacher education faculty because they were standardized, valid, and reliable. Each measure is discussed later in this article.

Methods

Setting and Participants

The PDP was created and implemented at a public university. The university is considered highly selective by *U.S. News and World Repor*t (2012). As a research-focused university, the majority of the 99 faculty members of the school of education are expected to engage in ongoing research projects.

The School of Education Teacher Education Program consists of two programs. One program is a five-year Bachelor of Arts plus Master's degree of Teaching (BAMT) that leads to initial teacher certification. Another program is a two-year Post-Graduate Master's of Teaching (PGMT), designed for individuals who elect to pursue a teaching license

Table 1
Measures in the Participant Data Pool

	Intro Course Students	3rd-Year BAMT	4th-Year BAMT 1st-Year PGMT	5th-Year BAMT 2nd-Year PGMT
Demographic Survey	X			
Common Core Survey				
• Ideas About Children (Schaefer & Edgerton, 1985)				
• Teacher Self-Efficacy Scale (Tschannen-Moran & Hoy, 2001)				
• NEO-5 Factor Inventory (Costa & McCrae, 1992)				
• Adult Attachment Scale (Griffin & Bartholomew, 1994)		X	X	X
• Depression Anxiety Stress Scale (Lovibond & Lovibond, 1995)				
• Teacher Multicultural Attitude Survey (Ponterotto, Baluch, Greig, & Rivera, 1998)				
• Culturally Responsive Teaching Self-Efficacy Scale (Siwatu, 2006)				
• Factors Influencing Teaching Choice (Watt & Richardson, 2007)				
Video Assessment of Interactions and Learning (Jamil, Sabol, Hamre, & Pianta, under review)	X		X	X
Classroom Assessment Scoring System (Pianta, Hamre, & La Paro, 2006)				X

Note. BAMT: Bachelor of Arts and Master's of Teaching combined 5-year program; PGMT: Post-Graduate Master's of Teaching 2-year program.

after completing an undergraduate degree. The two programs are integrated such that the BAMT and PGMT students take classes together. In total, the teacher education program graduates approximately 130 students each year.

The graduating class of 2012 represents a typical cohort for the teacher education program. The students are 80.2% female and 19.8% male. In terms of self-identified ethnicity, the composition of the student enrollment is 75% Caucasian, 12% Asian/Pacific Islander, 6% African American, 2% Hispanic, and 5% "unspecified." The number of students in each program is presented in Table 2. The largest number of students is found in the elementary education program (34%), while the second most popular programs are English education (15%) and social studies education (14%). The majority of teacher education students (61%) are in the five-year bachelor's plus master's degree. The largest teaching specialty is elementary (34%), followed by English education (15%) and social studies (14%).

Procedures

Although developed within the Teacher Education Department (TED), the PDP later moved out of the department and was administered by a research center affiliated with the School of Education. During its stay in the research center, the PDP took its current form, which is based largely on the data collection models of psychology departments. Mea-

Table 2
Class of 2012 Enrollment in Teacher Education Programs

Teaching Specialty	%
Early Childhood	4
Elementary	34
English	15
Foreign Language	7
Health/PE	3
Mathematics	6
Science	8
Social Studies	14
Special Education	9
BAMT/PGMT	**%**
BAMT	61
PGMT	39

Note. BAMT: Bachelor of Arts and Master's of Teaching combined 5-year program; PGMT: Post-Graduate Master's of Teaching 2-year program.

sures were selected by researchers in the affiliated research center and not by the faculty engaged in the teacher preparation program. In the 2009-2010 academic year, the responsibility of the administrating the PDP was returned to TED. These changes in who has overseen the PDP have determined not only its measures but also shaped the interactions that TED faculty have had with the PDP. Because the measures and the structure of the PDP were determined outside of the TED program, many TED faculty are not aware of what measures are included in the PDP or the specifics of the mechanics of the data collection system.

Currently, the PDP is administered by a doctoral student who is responsible for running the PDP website, working with faculty and graduate student researchers, and monitoring student research credits. The position is budgeted for 20 hours per week for the tasks related to the PDP.

Students are required to complete five research credits per academic year. The general PDP guidelines assume that one research credit equals one hour of participation in a research study or alternative assignment. Credits may be accrued throughout the academic year; however, students are held accountable for completing their credits in clinical courses during the spring semester. Therefore, the vast majority of research credits are offered and earned during the spring semester.

Student credits for completed studies as well as study sign-ups are managed through web-based software. Students are given an anonymous identification number that allows for data to be linked confidentially across different research studies. The software allows students to access online surveys and assessments as well as to register for times to participate in face-to-face research activities. Research activities available to students include those required by TED and participation in studies developed by other researchers either within the School of Education or outside researchers.

The PDP is also designed to allow researchers access to teacher education students either for research on teacher education or for research on other topics such as research projects focused on students in an undergraduate or graduate program. Researchers who want to use the PDP to collect participants must first obtain approval from the Institutional Review Board responsible for the protection of human subjects and then be approved by a committee that oversees the PDP. Once a study is approved, researchers can use the PDP website to recruit participants and collect data.

Instruments

To provide a context to discuss the many instruments included in the PDP, it is helpful to return to the guiding questions presented by

McNergney and Imig (2006). McNergney and Imig's questions were comprehensive, and each measure selected for use in the PDP addresses a piece of these questions. In this section, data from the most recent academic year, 2011-2012, are used to describe the psychometric properties of the measures.

What are the intellectual, physical, emotional, and social capacities of preservice teachers that are likely to influence their teaching behavior? It is not known which intellectual, physical, emotional, and social capacities of preservice teachers are likely to influence teaching behavior, although there is some preliminary evidence that a combination of these factors may predict teaching performance (Rockoff, Jacob, Kane, & Staiger, 2008). The PDP collects data that focus on students' personalities, emotional capacities, and social capacities. Taken together, these data can help to describe teacher education students at one university and to make connections between these traits and capacities and effective teaching or proclivity toward growth in instructional performance.

Personality measure. Are some people pre-disposed to be better teachers than are others? Answering this empirical question is important to teacher education. If a fairly stable characteristic, such as a teacher's personality type, predicts aspects of his or her success in the classroom, then teacher education programs would need to consider personality differences, either by admitting only those preservice teachers who would be successful or by scaffolding the program to meet the needs of individuals with a variety of personality types.

While there are several measures of personality, in recent decades, an empirical strategy for assessing personality types, known as the five-factor inventory, has become popular (Rockoff et al., 2008). At this university, preservice teacher personality is measured by the Neo Five-Factor Inventory (NEO-FFI), which identifies five personality factors: neuroticism, extraversion, openness, agreeableness, and conscientiousness (Costa, McCrae, & Dye, 1991). Participants respond to 60 items on a 5-point Likert scale (1=strongly disagree to 5=strongly agree), for which higher responses indicate an inclination for a certain personality type. Items from the different factors include, "I often feel inferior to others" (neuroticism, α=.869); "I like to have a lot of people around me" (extraversion, α=.826); "I often try new and foreign foods" (openness, α=.759); "Most people I know like me" (agreeableness, α=.834); and, "I keep my belongings neat and clean" (conscientiousness, α=.883).

Data from the PDP have shown that personality is largely stable for college students in the five-year bachelor's plus master's degree program (Ripski, LoCasale-Crouch, & Decker, 2011; Wiens & Ruday,

under review). Jamil, Downer, and Pianta (2012) also found a significant link between personality and preservice teacher self-efficacy. However, there are mixed results in regard to the association between personality and teaching performance in student teaching. Ripski et al. found that preservice teachers who reported higher levels of extraversion at program exit demonstrated lower levels of teaching quality. However, Wiens and Ruday found no significant association between personality and teaching performance at program exit but did find that personality predicted satisfaction with teaching.

Depression and anxiety measure. Unlike personality, which is more stable over time, emotional states are more fluid. Even though these traits may be more effective, it is still worthwhile to understand whether they have an impact on teaching performance. While depression and anxiety have been shown to impair work effectiveness in non-teaching settings (Haslam, Atkinson, & Brown, 2005), there is limited research on the relationship between these emotional states and factors related to teaching.

The Depression and Anxiety Stress Scale (DASS; Lovibond & Lovibond, 1995) is a 21-item self-report scale that asks respondents to describe aspects of their prior week's experiences. Responses are answered on a 4-point Likert scale of frequency: 1=did not apply, 2=applied some of the time, 3=applied a good part of the time, and 4=applied most of the time. Lovibond and Lovibond developed three composites factors: depression, anxiety, and stress. Items include, "I couldn't seem to experience any positive feelings at all" (depression, α=.893); "I felt scared without any good reason" (anxiety, α=.841); and, "I found it difficult to relax" (stress, α=.856).

Using data collected through the PDP, Ripski et al. (2011) found that depression was correlated with personality. Depression also predicted teaching performance in student teaching. Preservice teachers who reported higher levels of depression were observed to have lower quality teaching performance in their student teaching semester.

Attachment measure. Attachment theory (Ainsworth, Blehar, Water, & Wall, 1978; Bretherton, 1985) posits that, as people feel more comfortable and safe in the presence of an attachment figure, they are more willing to explore and take risks. Individuals with secure adult attachment may provide a better learning environment for children. As such, teachers who create a positive climate in their classrooms help students to learn more effectively (Hamre & Pianta, 2007).

The Adult Attachment Scale (AAS) is a modified version of the Relationship Scales Questionnaire (Griffin & Bartholomew, 1994). The original version of the Relationship Scales Questionnaire contains 30

short statements for which participants rate the extent to which each describes their characteristic style in close relationships. Because several of the questions are not germane to the study of preservice teachers, there was a need to develop a smaller, more specific scale for preservice teachers. A factor analysis was conducted on responses from the PDP sample, and, from this analysis, a single factor, Adult Attachment, was created. Responses are based on a 5-point scale, with the following anchors: 1=not at all like me, 3=somewhat like me, 5=very much like me. Questions include, "I find it easy to get emotionally close to others" and, "I know that others will be there when I need them." Internal consistency of the scale is quite high, α=.919.

One study has used the AAS data from the PDP. Ripski et al. (2011) found that adult attachment was correlated with the NEO-FFI and DASS. However, no predictive relationship was found between adult attachment and teaching performance during the student teaching semester.

Social capacity with culturally diverse students. Empirical research has shown that the forming of positive relationships between teachers and students is crucial for learning to occur (National Research Council, 2000). For positive relationships to develop, teachers must develop the skills to understand and to relate to a diverse group of students. The United States is undergoing a demographic change from a Caucasian majority nation to where people of color will represent the majority (Shrestha & Heisler, 2011); however, the majority of preservice teachers are Caucasian (Milner, Flowers, Moore, Moore, & Flowers, 2003). Assessing preservice teachers' racial attitudes may help teacher education programs to better understand these individuals' preparedness to teach.

The Teacher Multicultural Attitude Survey (TMAS) is a self-administered assessment of a participant's sensitivity to and awareness of multiculturalism (Ponterotto, Baluch, Greig, & Rivera, 1998). The TMAS consists of 20 questions that are answered on a 5-point Likert scale (1=strongly disagree to 5=strongly agree), with higher scores indicative of higher multicultural sensitivity and awareness. The scores for the 20 questions are combined to create one composite score for the measure. Questions in the TMAS include, "Teachers have the responsibility to be aware of their students' cultural backgrounds" and "As classrooms become more culturally diverse, the teacher's job becomes increasingly rewarding." The measure of internal consistency among PDP data for the TMAS score composite is well within the acceptable limits of reliability, at α=.857.

TMAS data have not been used as often by researchers as have the above-discussed measures. In the one study that used the TMAS data from the PDP, scores of 126 participants were examined for their atti-

tudes before and after the student teaching experience (Wiens, 2012a). In this study, participants' attitudes toward diversity were found to be stable (M=3.76) over the course of the study.

How do teachers in preparation demonstrate teaching behavior that has been shown to relate to or cause PK-12 student learning? The PDP uses the Classroom Assessment Scoring System (CLASS) as a measure of teaching effectiveness (Pianta, La Paro, & Hamre, 2008; Pianta, Hamre, & Mintz 2012). CLASS assesses teaching performance by providing a measure of teacher-student interactions (Pianta & Hamre, 2009; Pianta, La Paro, & Hamre, 2008). CLASS presents teacher-student interactions as organized into three domains: emotional support, classroom organization, and instructional support (Hamre, Pianta, Mashburn, & Downer, 2007). Each domain contains specific behaviors that have been identified as contributing to improved student learning. CLASS provides not only a framework for understanding important teaching behaviors but also a standardized measurement tool that enables the analysis of teacher-student interactions (Hamre & Pianta, 2007).

Research has shown CLASS to be a reliable and valid measure for grade levels (Cadima, Leal, & Burchinal, 2010; Graue, Rauscher, & Sherfinski, 2009; La Paro et al., 2009; Parkarinen et al., 2010). The "Measures of Effective Teaching" study used CLASS as one of its standardized observation measures and found that CLASS correlated with student achievement gains as measured by achievement tests (Gates Foundation, 2012).

In the fall semester of their final year, preservice teachers complete a one-semester student teaching placement. The preservice teachers video-record themselves during a period of time when they have taken on full teaching responsibilities. From the videos, trained raters generate two sets of CLASS codes that are then combined into one mean score. Raters were initially trained to reliability on the tool through a rigorous two-day training session, during which they learned the CLASS framework and conducted multiple practice tests. Next, raters passed a reliability test by demonstrating that they could use the CLASS tool successfully across multiple classroom situations. To be considered reliable, all raters were required to pass a reliability test for which their coding needed to match within a margin of one point to a master coding list 80% of the time.

PDP CLASS data has been used in multiple studies as a measure of observed teaching performance (Jamil, Downer, & Pianta, 2012; Ripski et al., 2011; Wiens, 2013). CLASS coding provides a differentiation between low quality (coded 1-2), moderate quality (coded 3-5), and high quality (6-7). Jamil et al. examined data from 509 preservice teachers and found that all three dimensions of teaching performance fell in the moderate quality

category (emotional support, M=5.52; classroom organization, M=5.06; instructional support, M=4.26). Wiens, in a study of a different set of 176 preservice teachers found similar results (emotional support, M=5.33; classroom organization, M=4.91; instructional support, M=3.74).

How do preservice teachers' students think about teaching and learning—are they motivated, challenged, and reinforced by teachers' actions? Preservice teachers enter their teacher education programs with preconceived ideas about teaching and learning. Understanding how these ideas affect how they make sense of the teacher education program is an important part of crafting programs that meet the needs of preservice teachers. The PDP is used to examine different attitudes and knowledge that preservice teacher possess. As a longitudinal database, it also enables the analysis of how these attitudes and knowledge change over time.

Attitudes about children. Contemporary educational thinking has moved away from favoring the authoritarian role of teaching toward a focus on student needs. This evolution of thought can be seen in the latest version of the Model Core Teaching Standards (MCTS), which begins with an emphasis on the learner and learning (Council of Chief State School Officers, 2010). According to the MCTS, teachers need to work with learners to support individual student needs. There is a clear emphasis in the standards on teachers' constructing a more child-centered, effective learning environment.

How preservice teachers conceptualize learning and instruction may be connected to the way that they perceive children. Preservice teachers' attitudes about children were measured using the Modernity Scale (Schaefer & Edgerton, 1985). The Modernity Scale differentiates between more traditional adult-centered, or authoritarian, views and more progressive, child-centered views. The scale consists of 16 items that ask participants to respond on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). All items in the Modernity Scale are combined into one factor (α=.741), for which larger numbers indicate more child-centered views.

Schaefer and Edgerton (1985) developed the Modernity Scale as a measure of parental attitudes toward child rearing. More recently, however, the Modernity Scale has been used by researchers to understand teachers' attitudes toward children (see Burchinal, Peisner-Feinberg, Pianta, & Howes, 2002; Mashburn, Hamre, Downer, & Pianta, 2008; Scott-Little, La Paro, & Weisner, 2006).

Drawing on the PDP data, the Modernity Scale also has been used to examine preservice teachers' attitudes toward children. The analysis of the PDP data demonstrated that more child-centered or progressive beliefs were associated with higher preservice teacher self-efficacy (Jamil

et al., 2012). There is also evidence that preservice teachers' ideas about children become more child-centered as they progress through the teacher education program (Wiens, Hessberg, & Baran, 2012).

Measures of teacher self-efficacy. Teacher self-efficacy, or the confidence of an individual to be successful at a specific task, has been linked to many positive outcomes for teachers. Teachers with high self-reported self-efficacy are more satisfied with their careers (Coladarci, 1992) and are more likely to persist in teaching (Burley, Hall, Villeme, & Brockmeier, 1991). Teacher self-efficacy and resiliency have been linked (Yost, 2006), and, thus, measurement of preservice teacher self-efficacy is important.

The PDP contains data from two measures of teacher self-efficacy. The first measure is the Teacher Self-Efficacy Scale (TSE; Tschannen-Moran & Hoy, 2001). The TSE scale consists of 24 items, comprising three composite subscales. The efficacy for instructional strategies composite (α=.905) is typified by questions such as, "How well can you respond to difficult questions from your students?" An example of the efficacy for classroom management composite (α=.913) is, "How much can you do to control disruptive behavior in the classroom?" Finally, efficacy for student engagement (α=.896) is characterized by questions such as, "How much can you do to help your students value learning?" Responses are based on a 9-point Likert scale with the following anchors: 1=nothing, 3=very little, 5=some influence, 7=quite a bit, and 9=a great deal.

As noted, Jamil et al. (2012) used the TSE scale data and found a connection between self-efficacy and personality types. Jamil et al. found that preservice teachers who were more outgoing had higher self-efficacy at program exit. Additionally, teachers with more child-centered views had a higher teaching self-efficacy. Jamil et al. did not, however, find a significant relationship between self-efficacy and teaching performance, as measured by the CLASS instrument.

The second measure of self-efficacy, the Culturally Responsive Teaching Self-Efficacy Scale (CRTSE; Siwatu, 2006) concerns preservice teachers' confidence in their ability to work with culturally diverse learners. The CRTSE measures participants' culturally responsive teaching self-efficacy and outcome expectancy beliefs and is a validated measure (see Castro, 2010; Chu, 2011). The CRTSE asks participants to rate their confidence, on a scale of 0 to 100, in accomplishing a list of 41 tasks. Three points in the scale are as follows: 0=no confidence at all, 50=moderately confident, and 100=completely confident.

The 41 items in the CRTSE are combined into a single score (α=.586) that represents the participant's self-efficacy in regard to his or her ability to meet the needs of diverse learners. Higher scores indicate that

the participant has greater confidence in his or her abilities. Examples of items in the CRTSE include, "I am able to use a variety of teaching methods" and "I am able to develop a community of learners when my class consists of students from diverse backgrounds."

CRTSE data have been used in one study thus far. As noted, Wiens (2012a), in a study discussed above, examined multicultural teaching self-efficacy before and after a one-semester student teaching experience. The study found that preservice teachers displayed a significantly higher level of self-efficacy in this domain following student teaching.

Understanding the reasons for becoming a teacher and satisfaction with that choice. An understanding why individuals choose to go into teaching and how satisfied they are with this decision can be helpful for teacher education programs in recruiting and training future teachers. Examining the reasons for teaching and the satisfaction with that choice provide information on which preservice teachers may be more likely to be committed to the profession over the long term. Research indicates that teachers who enter the profession with a high level of commitment to teaching tend to stay in the profession longer than do those with a lower level of commitment (Chapman & Green, 1986).

The Factors Influencing Teaching Choice (FIT-CHOICE) scale measures factors that influence preservice teachers' choice to teach and their feelings about the decision to become a teacher (Watt & Richardson, 2007). The FIT-CHOICE scale demonstrates strong validity and reliability in teacher education settings (e.g., Eren & Tezel, 2010; Lawver & Torres, 2011). School of Education researchers selected one set of 13 items from the full FIT-CHOICE measure because they capture most of the factors included in the full measure. A second set of 5 items concerns feelings about the choice of a teaching career and likewise capture some of the factors included in the full FIT-CHOICE.

The first section of the FIT-CHOICE begins with the prompt, "I chose to become a teacher because . . . ," followed by 13 items that participants rate on a 7-point Likert scale (1=not at all important to 7=extremely important). The second section begins with the prompt, "Your thoughts regarding teaching." Participants rate their responses to the statements on the same 7-point Likert scale. Each of the items selected for use in the Common Core Survey (CCS) was adapted from a different factor in the original measure; therefore, no composites were created from the FIT-CHOICE.

Currently, only one study (Wiens & Ruday, under review) has used the PDP FIT-CHOICE data. In this study, Wiens and Ruday looked at the relationship between personality and satisfaction with teaching as

a career choice. The results indicated that personality was related to satisfaction with teaching. Specifically, those individuals who have a tendency toward neuroticism are generally less satisfied with teaching, plan to put less effort into teaching, and do not plan to persist in the profession. Extraverted individuals, in comparison, showed a higher satisfaction with teaching, planned to persist in the profession, and planned to seek leadership roles.

Measuring students' knowledge of effective teaching. There is a lack of standardized measures of preservice teacher learning that can be used across programs and settings. To address this problem, TED implemented the Video Assessment of Interactions and Learning (VAIL), which was developed by researchers to examine teachers' ability to identify effective teaching strategies and behaviors in videos of classroom teachers (Hamre, et al., 2012; Jamil, Sabol, Hamre, & Pianta, under review). VAIL is based on the dimensions of CLASS, and videos were selected for their demonstration of effective teaching based on CLASS evaluations. Participants watch three short videos of in-service pre-kindergarten language arts classes and identify up to five strategies and five specific behavioral examples of those strategies from the videos. Responses are then coded for accuracy of identification of the strategies and examples. Coding also indicates whether the examples match the strategies. The coding results are then combined across the three videos into a summary score that is used for analysis (Wiens, Hessberg, LoCasale-Crouch, & DeCoster, 2013; Jamil et al., under review).

VAIL data, drawn from the PDP, have been shown to be related to observed teaching performance, as measured by CLASS (Wiens, 2013). The data also have shown that VAIL can be reliably implemented at multiple points in a teacher education program (Wiens et al., 2013) and that it is not biased toward any particular teacher education specialty (Wiens 2013; Wiens & Hessberg, 2011; Wiens et al., 2013).

Results: Use of the Participant Data Pool

In this article, the author sought to understand the feasibility of collecting data on teacher education programs that could be used for evaluation. In this regard, I will first discuss participation in the PDP and how implementation of the PDP has been achieved. Then I will examiner researchers' use of the PDP.

Student Participation in the PDP

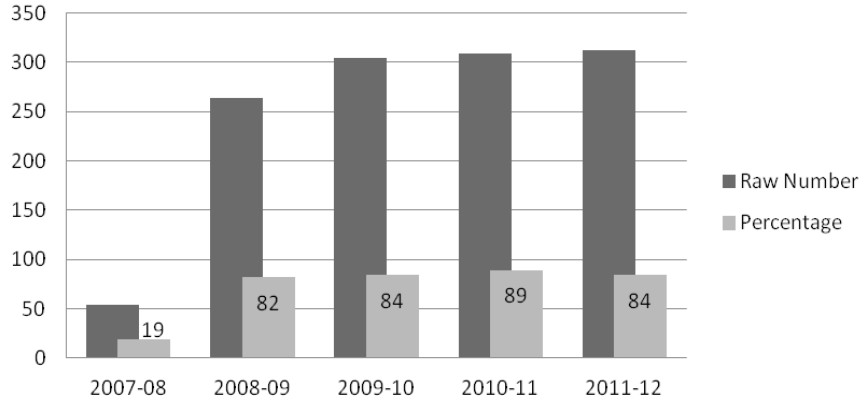
To understand the feasibility of collecting and analyzing data on teacher education students using the PDP, it was helpful to begin by

examining preservice teacher participation in the effort. The central focus of the PDP data collection has been the CCS, which has been a requirement of all teacher education students in each year of their program, beginning in the 2008-2009 academic year. Data on participation rates for the CCS are illustrated in Figure 2 The dark bars indicate the raw number of students who are expected to participate in the CCS, while the lighter line shows the percentage of all teacher education students who participated. Participation in the CCS has increased in terms of raw numbers over the five years examined in this paper. However, the percentages of participation have remained fairly consistent (82%-89%) since the CCS became a requirement for most students in the teacher education program. Even though students could elect to complete an alternative assignment, participation has reminded high and reasonably consistent.

A certain percentage of students do not participate in the CCS for various reasons. Students may not participate because they are not aware of the requirement or because they simply choose not complete it. Students who take their education courses out of order or take a semester off also might not participate. Additionally, certain programs within the larger teacher education program promote participation in the CCS more than do others. Therefore, some students are not encouraged to participate to the same degree that others are.

As noted earlier, another initiative of the teacher education program is to collect CLASS data on each preservice teacher; however, the achievement of this goal varies from year to year. As illustrated in Figure 3,

Figure 2
Participation in the Common Core Survey



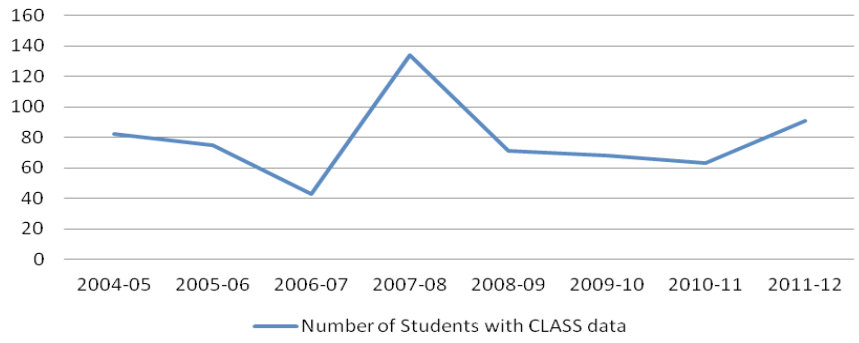
with a few exceptions, TED has successfully collected CLASS data on approximately half the students each year.

Because CLASS is a standardized observational protocol, data are on preservice teachers more difficult to collect as compared to data from a survey or administrative data. CLASS raters need to be trained to reliability in the measure to provide useful data for the teacher education department. The department typically had trained the university supervisors in CLASS coding, and these individuals have coded their own students' performance. However, not all university supervisors have achieved reliability, and the student teachers of these supervisors may not get evaluated using CLASS. The more recent use of videos to code student teaching performance has the potential to lead to increased numbers of data on student teachers.

Even though the PDP contains CLASS data on approximately half of the student teachers, over eight years it has amassed complete CLASS data on nearly 630 student teachers. The collected CLASS scores can provide picture of the average teaching performance of the school's student teachers over an eight-year period. As a standardized observation measure, CLASS requires training and resources to implement; however, the teacher education program has had some success in collecting data on its student teachers.

The final data collection initiative of the PDP is VAIL. Much like CCS, VAIL is administered online, which provides students with the opportunity to complete it at their leisure. VAIL is required of students in the introductory course, which is a prerequisite for admission into the program. These students may or may not apply for admission into the teacher education program. Students in the introductory course complete VAIL in the first three weeks of the semester. The students who do en-

Figure 3
CLASS Data Collected on Student Teachers



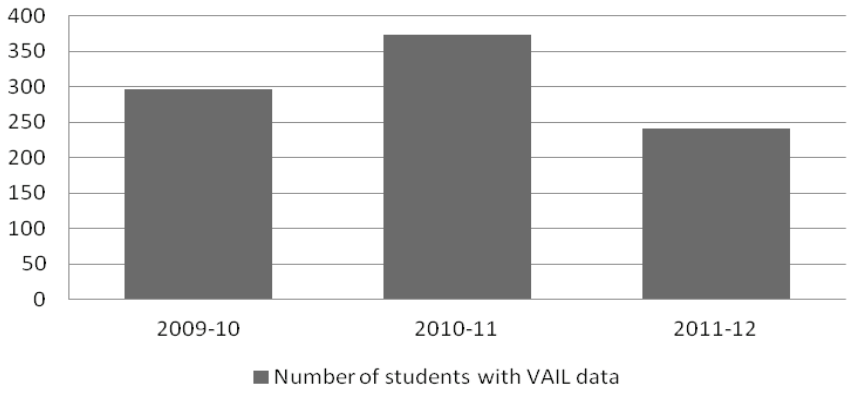
roll in the five-year bachelor's plus master's teacher education program will then take the VAIL again in the semester prior to student teaching in their fourth year and again following student teaching in their fifth year. Students in the post-graduate master's of teaching program take the VAIL in their second semester prior to student teaching and again in their final semester following student teaching.

VAIL responses are collected through the PDP and coded by trained raters. These raters are doctoral students who are paid a summer stipend to code the data. The raters are trained at the end of the spring semester and code the data over the summer months. Participation in VAIL varies, depending on the number of students enrolled in the introductory course each year (Figure 4). In each of the three years for which VAIL data were collected, more than 200 students took the assessment. As opposed to CLASS and the CCS data, VAIL data can potentially show growth in teacher education students' abilities to understand and recognize effective teaching.

An Examination of Researcher Use of the PDP

A key underlying question of data collection and management in the PDP is to what extent the system can be used to contribute to the understanding of teacher education students (McNergney & Imig, 2006). The PDP was designed not only as a data collection and management system for the teacher education department, but also as a means for researchers to access teacher education students. This has led the PDP to be used by researchers who desired to understand teacher education

Figure 4
Number of Students with VAIL Data

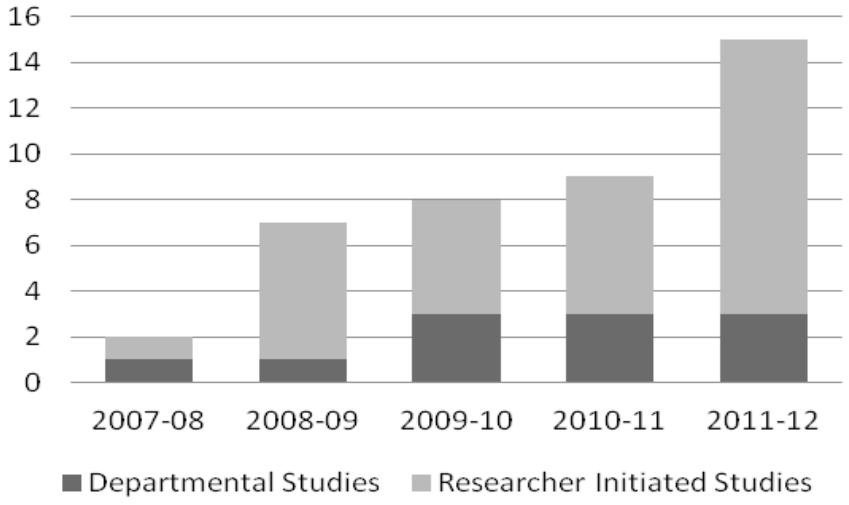


students and related issues and researchers who are more broadly interested in the young adult population.

Figure 5 presents the number of studies that used the PDP to collect data both directly through online surveys and indirectly through its function as a scheduling device for face-to-face data-gathering efforts. Included in the figure are teacher education department-sponsored studies, such as CCS and VAIL, and researcher-initiated studies. While the number of studies developed by the teacher education program has remained fairly stable over the five-year period of 2007-2008 to 2011-2012, the use of the PDP by researchers for data collection has grown steadily from one researcher-initiated study in the 2007-2008 academic year to 13 in the 2011-2012 academic year.

In the five years that the PDP has operated in its current format, there have been a total of 25 studies, some of which have spanned multiple years. Figure 6 presents the breakdown of who have initiated these studies. As shown in the figure, the majority of studies in the PDP have been created by faculty. When combining TED faculty and other university faculty, this accounts for 64% of the studies in the PDP. In total, TED faculty, including the TED required measures (i.e., CCS, CLASS, and VAIL), accounted for nine studies, other faculty created seven studies, doctoral students were responsible for eight studies, and the authorship of one study could not be determined.

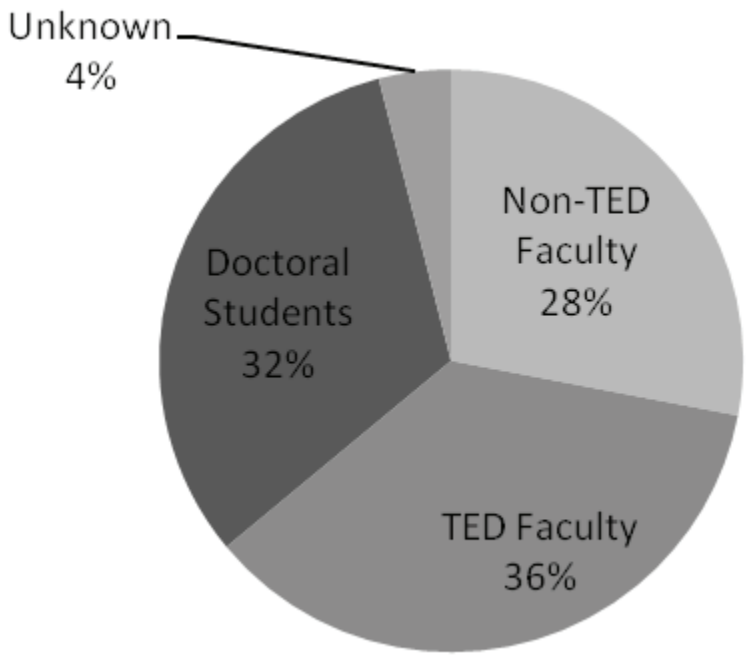
Figure 5
Research Studies in the PDP



In the last year, a new type of study has emerged from the PDP. Previously, the studies associated with the PDP concerned teacher education students. As more faculty have become aware of the PDP, however, researchers have begun to use the PDP as a way to recruit college-aged participants for studies. Two of the recent studies have used the PDP as a means for recruiting a control group of college students for an intervention used on a non-teacher education group. Likewise, another study drew upon the PDP for recruiting college-aged students. This particular study tapped into the PDP to recruit a control group of university-aged students but did not focus on issues specific to education.

Generally, studies that have drawn upon the PDP have been part of a dissertation or have as their goal publication in peer-reviewed journals. Nevertheless, two of the studies that used the PDP were pilot studies, with Institutional Review Board approval, but were not directly used for publications. The pilot studies recruited small numbers of participants to test an instrument or interview protocol prior to implementing it in

Figure 6
PDP Study Indicators



a research study. Another three studies were for internal use only and were not intended for dissemination.

Use of Data by Teacher Education Department

In addition to the use of PDP data for research, the data also were intended to be used for program development. Only recently, however, has TED begun to examine PDP data for the information it provides in regard to program success and future development. As a result of undergoing a national accreditation process, the department has begun to examine specific measures from the PDP, including CLASS, VAIL, the Modernity Scale, TSE, CRTSE, and TMAS. These measures have helped TED to examine differences in the data between students in the different TED programs such as special education, elementary education, and secondary education. \The knowledge gained by the use of these measures can be used to change the program to better meet the needs of preservice teachers.

Limitations

There are certain limitations to the way the PDP was constructed and is currently implemented. As noted, much of the PDP was designed by researchers outside TED. This has led to a lack of large-scale participation by teacher education faculty and a lack of understanding of the measures and purpose of the PDP. Of the researcher-initiated studies conducted using the PDP, only about one-third have been conducted by teacher education faculty. Additionally, although there is a wealth of data collected that can be drawn from the PDP, there is little attempt by faculty to use the data to improve the program or to understand students better.

The lack of teacher education faculty participation also has contributed to the lack of full participation in the centerpiece of the PDP, the CCS. While the goal of the PDP is to collect data on every student through the CCS, this has not been accomplished. Some professors are stricter than others about enforcing the PDP participation requirement in their courses. Additionally, one small teacher education program has taken itself out of participation in the PDP because the program has developed its own research requirements. Had the PDP been created by faculty within the department, perhaps it would be more meaningful to that faculty, which would encourage higher participation rates.

Generally, the PDP has been successful in collecting data on the CCS. Participation rates continue to be high but not as high as TED would like to see. As noted, there is a small program that does not fully participate

in the PDP, which accounts for some of the missing participants. The main limitation in PDP data collection, however, is in how students are held accountable. Without some form of accountability, participation rates would be significantly lower. The PDP holds students accountable through clinical courses, but not all students take the courses in the same order. There is a small group of students who take a year or semester off or study abroad, and these students could be missed by the current accountability structure. Previously, the PDP had been tracked through a one-credit course in addition to the other courses that students were required to take. This was less effective compared to the current system because some students either did not register for the class or they had problems registering for the course because they already were taking a full load of classes, which prevented them from registering without paying an additional fee. The current system has been an improvement.

The most difficult data to collect in the PDP has been student teaching performance data drawn from CLASS. The rates for collecting CLASS data on student teachers have been highly inconsistent, ranging from only one-third to nearly all student teachers. TED has managed to save money by training university supervisors in CLASS and having them code their own student teachers' performance. This can lead to problems, however, when university supervisors either do not achieve reliability in CLASS or simply do not provide CLASS evaluations on their student teachers to the TED office. Stricter oversight by the TED office may help improve the amount of CLASS data collected each year. Further, to ensure that each student teacher has CLASS evaluations, TED could hire trained coders to evaluate the student teaching videos.

Discussion

The university personnel who came together to create the PDP because they believed it was feasible to create and manage data on teacher education students (McNergney & Imig, 2006). Initially, the PDP was created to address the TNE Project's main goal of creating outstanding teacher education programs (Carnegie Corporation of New York, 2001). The university's experience with the PDP largely validated the personnel who originally created it.

This article has shown that it is feasible to create and manage teacher education data in a national research university. While the PDP took several years to get off the ground, it has gained traction at the university. Use of the PDP by researchers has steadily grown over the past five years. Further, a research requirement, along with increased attention to the PDP, has led to increasingly high percentages of student

participation in research activities. This effort has led to a database of information on more than 1,300 students over a five-year period.

The experiences of this university's data gathering effort also demonstrate the need for universities to commit resources toward the collection and analysis of teacher education data. The university commits a full-time doctoral student to the management of the PDP and provides the student with a stipend, health insurance, and tuition remission. Additionally, TED bears the expenses of the department-initiated research projects, including the fees for the software needed to collect and manage the data.

When a school of education commits resources and demands so much of its students' time to participating in the PDP, it is important to understand whether the PDP is worthwhile in terms of its contribution to the knowledge base in teacher education. Data drawn from the PDP have been the basis of several dissertations and empirical research papers, with more forthcoming. A handful of papers over a decade, however, is not a great dissemination of knowledge. The growth in use of the PDP by faculty researchers, however, provides hope that the knowledge to be gleaned from the PDP will increase in scope and in its dissemination.

The central goal of the TNE Project and, thereby, the underlying principle of the PDP was to "stimulate construction of excellent teacher education programs at selected colleges and universities" (Carnegie Corporation of New York, 2001, p. 1). Perhaps it is more appropriate to focus on whether the PDP has improved teacher education at its own university. A collection of longitudinal data on teacher education students allows the teacher education program to evaluate growth in students and to identify areas of the program that should be examined. Development of the PDP is an ongoing process, and, so far, the PDP has been successful in systematically collecting data and making that data available to university researchers and analysts within the teacher education program.

Nevertheless, this article noted the unequal use of the data for research purposes. While researchers have conducted multiple analyses on some measures in the PDP, other measures have not been used frequently for analysis. As data are continually collected through the PDP, further analysis can provide information on how teacher education students grow personally and how their conceptions of teaching change throughout the course of their program. It is equally important to continually assess the usefulness of each measure in the PDP while searching for other measures that might be incorporated into the data collection effort.

The PDP and the knowledge it produces are necessarily constrained by the context of the university in which it exists. Data from the PDP can inform teacher education practice, but generalizations to other programs are difficult, given the highly selective nature of the institution.

Further, the PDP has not extended its reach to program graduates and their performance as fully certified in-service classroom teachers. Future research should connect data from teacher education students to their performance once they are hired by schools. Collecting data that link preservice characteristics to in-service performance is the next logical step to gaining better understanding of teacher education students and teacher education programs.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillside, NJ: Lawrence Erlbaum Associates.

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., & Sheppard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers* (EPI Briefing Paper 278). Washington, DC: Economic Policy Institute.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2005). *How changes in entry requirements alter the teacher workforce and affect student achievement* (NBER Working Paper No. 11844). Cambridge, MA: National Bureau of Economic Research.

Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *The Future of Children, 17*(1), 45-68.

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*, 416-440.

Bretherton, I. (1985). Attachment theory: Retrospect and prospect. *Monographs of the Society for Research in Child Development, 50*(1/2), 3-35.

Brophy, J. E. (1999). *Teaching* (Educational Practices Series-1). Geneva, Switzerland: International Academy of Education and International Bureau of Education, UNESCO.

Burchinal, M. R., Peisner-Feinberg, E., Pianta, R., & Howes, C. (2002). Development of academic skills from preschool through second grade: Family and classroom predictors of developmental trajectories. *Journal of School Psychology 40*(5), 415-436.

Burley, W., Hall, B., Villeme, M., & Brockmeier, L. (1991). A path analysis of the mediating role of efficacy in first-year teachers' experiences, reactions, and plans. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Buzzelli, C. & Johnston, B. (2001). Authority, power, and morality in classroom discourse. *Teaching and Teacher Education, 17*(8), 873-884.

Cadima, J., Leal, T., & Burchinal, M. (2010). The quality of teacher-student interactions: Associations with first graders' academic and behavioral outcomes. *Journal of School Psychology, 48*, 457-482.

Carnegie Corporation of New York. (2001). *Teachers for a new era: A national initiative to improve the quality of teaching*. New York: Author.

Castro, A. J. (2010). Themes in the research on preservice teachers' views of cultural diversity: Implications for researching millennial preservice teachers. *Educational Researcher 39*(3), 198-210.

Chapman, D. W., & Green, M. S. (1986). Teacher retention: A further examination. *Journal of Educational Research, 79*, 273-279.

Chu, S.-Y. (2011). Teacher efficacy beliefs toward serving culturally and linguistically diverse students in special education: Implications of a pilot study. *Education and Urban Society*. Retrieved from http://eus.sagepub.com/content/early/2011/05/26/0013124511409401

Cochran-Smith, M., & Zeichner, K. M. (2005). *Studying teacher education: The report of the AERA Panel on Research and Teacher Education* [Executive summary]. Mahwah, NJ: Lawrence Erlbaum Associates.

Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education, 60*, 323-337.

Constantine, J., Player, D., Silva, T., Hallgren, K, Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification*, Final Report (NCEE 2009-4043). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Council of Chief State School Officers. (2010). *Model core teaching standards: A resource for state dialogue*. Washington, DC: Council of Chief State School Officers Interstate Teacher Assessment and Support Consortium.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO- PI-R) and NEO five-factor inventory (NEO-FFI) professional manual*. Odesssa, FL: Psychological Assessment Resources.

Darling-Hammond, L. (2006a). Constructing 21st-century teacher education. *Journal of Teacher Education, 57*, 300-314.

Darling-Hammond, L. (2006b). Assessing teacher education: The usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education, 57*, 120-138.

Eren, A., & Tezel, K.V. (2010). Factors influencing teaching choice, professional plans about teaching and future time perspective: A meditational analysis. *Teaching and Teacher Education, 26*, 1416-1428.

Floden, R. E. (2008). Improving methods for research on teacher education. In M. Cochran-Smith, S. Feiman-Nemser, D. J. MacIntyre, & K. E. Demers (Eds.), *Handbook of research on teacher education* (3rd ed., pp. 1183-1188). New York: Routledge.

Gates Foundation (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Author.

Glazerman, S., Loeb, S., Goldhaber, D., Staiger, D., Raudenbush, S., & Whitehurst, G. (2010). *Evaluating teachers: The important role of value-added*. Washington, DC: Brown Center on Education Policy at Brookings.

Goldhaber, D. (2008). Teachers matter; effective teacher quality policies are elusive. In H. F. Ladd & E. B. Fiske (Eds.). *Handbook of research in education finance and policy* (pp. 146-165). New York: Routledge.

Graue, E., Rauscher, E., & Sherfinski, M. (2009). The synergy of class size reduction

and classroom quality. *The Elementary School Journal, 110*(2), 178-201.
Greenberg, J., McKee, A., & Walsh, K. (2013). *Teacher prep review: A review of the nation's teacher preparation programs*. Washington, DC: National Council on Teacher Quality.
Griffin, D., & Bartholomew, K. (1994). Models of the self and other: Fundamental dimensions underlying measures of adult attachment. *Journal of Personality and Social Psychology, 67*(3) 430-445.
Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal, 45*, 184-205.
Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in pre-school and early elementary classrooms. In R. Pianta, M. Cox, & K. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49-84). Baltimore, MD: Brookes.
Hamre, B. K., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2007). *Building a science of classrooms: Application of the CLASS framework in over 4,000 U.S. early childhood and elementary classrooms*. New York: Foundation for Child Development. Retrieved from http://www.fcd-us. org/resources/resources_show.htm?doc_id=507559
Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., & Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge, and observed practice. American *Educational Research Journal, 49*(1), 88-123.
Haslam, C., Atkinson, S., Brown, S., & Haslam, R. (2005). Effects of medication for depression and anxiety on safety in the workplace. *Occupational and Environmental Medicine, 62*, 538-545.
Jamil, F. M., Downer, J. T., & Pianta, R. C. (2012). Association of pre-service teachers' performance, personality, and beliefs with teacher self-efficacy at program completion. *Teacher Education Quarterly, 39*(4), 119-138.
Jamil, F. M., Sabol, T. J., Hamre, B. K., & Pianta, R. C. (under review). Assessing teachers' skills in detecting and identifying effective interactions in the classroom: Theory and measurement. *Elementary School Journal*.
Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2006). *What does certification tell us about teacher effectiveness? Evidence from New York City*. (Working Paper 12155). Cambridge, MA: National Bureau of Economic Research.
La Paro, K. M., Hamre, B. K., Locasale-Crouch, J., Pianta, R. C., Bryant, D., Early, D., & Burchinal, M. (2009). Quality in kindergarten classrooms: Observational evidence for the need to increase children's learning opportunities in early education classrooms. *Early Education and Development, 29*(4), 657-692.
Lawver, R. G., & Torres, R. M. (2011). Determinants of pre-service students' choice to teach secondary agricultural education. *Journal of Agricultural Education, 52*(1), 61-71.
Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales*. Sydney, Australia: Psychology Foundation.
Mashburn, A. J., Hamre, B. K., Downer, J. T., & Pianta, R. C. (2006). Teacher and classroom characteristics associated with teachers' ratings of prekin-

dergartners' relationships and behaviors. *Journal of Psychoeducational Assessment, 24*(4), 367-380.

McNergney, R. F., & Imig, S. R. (2006). Challenges of assessing valued added to PK-12 education by teacher education. *Teachers College Record*. Retrieved from http://www.tcrecord/org/Content.asp?ContentID=12814

Medina, J. (2009, October 23). Teacher training termed mediocre. *The New York Times*, pp. A3.

Milner, H. R., Flowers, L. A., Moore, E. Jr., Moore, J. L. III, & Flowers, T. A. (2003). Preservice teachers' awareness of multiculturalism and diversity. *The High School Journal, 87*(1), 63-70.

Mitzel, H. E. (1960). Teacher effectiveness. In C. W. Harris (Ed.), *Encyclopedia of educational research* (3rd ed., pp. 1481-1486). New York: Macmillan.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

National Council for Accreditation of Teacher Education. (2008). *Professional standards for the accreditation of teacher preparation institutions*. Washington, DC: Author.

National Council for Accreditation of Teacher Education. (2011, May 12). *Guidelines on assessment*. Retrieved from http://ncate.org/Accreditation/ProgramReview/GuidelinesAndProcedures/GuidelinesonAssessment/tabid/446/Default.aspx

Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*, 237-257.

Parkarinen, E., Kiuru, N., Lerkkanen, M.-K., Poikkeus, A.-M., Siekkinen, M., & Numi, J.-E. (2010). Classroom organization and teacher stress predict learning motivation in kindergarten children. *European Journal of Psychology of Education, 25*(3), 281-300.

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.

Pianta, R. C., Hamre, B. K., & Mintz, S. (2012) *Classroom Assessment Scoring System (CLASS): Secondary manual.* Charlottesville, VA: Teachstone.

Pianta, R. C., La Paro, K., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS)*. Baltimore, MD: Paul H. Brookes.

Ponterotto, J. G., Baluch, S. B., Greig, T., & Rivera, L. (1998). Development and initial score validation of the Teacher Multicultural Attitude Survey. *Educational and Psychological Measurement, 58*, 1002-1016.

Ripski, M. B., LoCasale-Crouch, J., & Decker, L. (2011, Spring). Pre-service teachers: Dispositional traits, emotional states, and quality of teacher-student interactions. *Teacher Education Quarterly*, 77-96.

Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2008). *Can you recognize an effective teacher when you recruit one?* (Working Paper 14485). Cambridge, MA: National Bureau of Economic Research.

Sanders, W., & Horn, S. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 13*(3), 247-256.

Schaefer, E. S., & Edgerton, M. (1985). Parent and child correlates of parental modernity. In I. E. Sigel (Ed.), *Parental belief systems* (pp. 287-318). Hillsdale, NJ: Lawrence Earlbaum Associates.

Scott-Little, C., La Paro, K. M., & Weisner, A. (2006). Examining differences in students' beliefs and attitudes: An important element of performance-based assessment systems for teacher preparation programs. *Journal of Early Childhood Teacher Education, 27*, 379-390.

Shrestha, L. B., & Heisler, E. J. (2011). *The changing demographic profile of the United States*. Washington, DC: Congressional Research Service.

Siwatu, K. O. (2006). Preservice teachers' culturally responsive teaching self-efficacy and outcome expectancy beliefs. *Teaching and Teacher Education. 23*, 1086-1101.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.

U.S. Department of Education (2009, November). *Race to the top* [Executive summary]. Washington, DC: Author.

U.S. News and World Report. (2012). *National university rankings*. Retrieved from http://colleges.usnews.rankingsandreviews.com/best-colleges

Watt, H. M., & Richardson, P. W. (2007). Motivational factors influencing teaching as a career choice: Development and validation of the FIT-Choice scale. *The Journal of Experimental Education, 75*(3), 167-202.

Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2002) Teacher preparation research: An insider's view from the outside. *Journal of Teacher Education 53*, 190-204.

Wiens, P. D. (2012a, February). Student teaching and preservice teachers' attitudes towards working with diverse learners. Paper presented at the Curry Research Conference, Charlottesville, VA.

Wiens, P. D. (2012b). The missing link: Research on teacher education. *Action in Teacher Education, 34*(3), 249-261.

Wiens, P. D. (2013). Are they learning? Measuring preservice teachers' skill at detecting effective teaching interactions (Unpublished doctoral dissertation). University of Virginia, Charlottesville, VA.

Wiens, P. D., & Hessberg, K. (2011, December). Do you see what I see? Using video-based assessment to measure preservice teachers' ability to recognize effective teaching. Paper presented at the College and University Faculty of the National Council for the Social Sciences Annual Conference, Washington, DC.

Wiens, P. D., Hessberg, K., & Baran, J. (2012). Changes in attitudes towards children, personality, and future social studies teachers: An analysis of survey data. Paper presented at the College and University Faculty of the National Council for the Social Sciences Annual Conference, Seattle, WA.

Wiens, P. D., Hessberg, K., LoCasale-Crouch, J, & DeCoster, J. (2013). Using a standardized video-based assessment in a university teacher education program to examine preservice teachers knowledge related to effective teaching. *Teaching and Teacher Education, 33*, 24-33.

Wiens P. D., & Ruday, S. (in press). Personality and preservice teachers: Does it change, does it matter? *Issues in Teacher Education*.

Yost, D. S. (2006). Reflection and self-efficacy: Enhancing the retention of qualified teachers from a teacher education perspective. *Teacher Education Quarterly, 33*(4), 59-76.