The Sinking State of Public Colleges of Education: Lessons Learned Without Lifeboats

Audrey Amrein-Beardsley
Joshua H. Barnett
Mary Lou Fulton Teachers College, Arizona State University

ABSTRACT: The bow of the flagship is beginning to tip. Unless teacher educators can prove that their graduates are effective classroom teachers, they will not stay afloat in the ocean of accountability that has already encumbered U.S. public schools. In this article one consortium of teacher education leaders' and researchers' nascent efforts to research the worth of the state's public colleges of education are presented. Described are their journey, the distance covered, and the icebergs, so to speak, they have hit. Due to the vastness of this statewide effort, the lessons learned may help inform others as they too embark on similar journeys.

Given survivors' accounts, historians are certain that on the night of April 14, 1912, the *Titanic* sank to the sounds of eight of its second-class traveling band members. The band members played separately during only high-class occasions until this tragic night. They *went down* in history as a cohesive unit when their well-respected band leader called upon them to perform for the first time. Survivors not only noted how odd it was to be wearing life jackets and awaiting the preparation of their lifeboats while music was played, but they also noted that when the first class lounge emptied, the band members reassembled outside near the boat deck. The band continued to fill the crisp air with their melodies in accord and performed until it was humanly impossible to play on. Their memorial plaque in Liverpool, England, reads, "Courage and compassion joined, make the

hero and the man complete" (Scarth, 2009, p. 199).

This same courage and compassion language is often found describing public school teachers in the United States—their quixotic and resolute nature attempting to rectify any and all social ills presented in their classrooms. However, unlike the oft-lauded band of the *Titanic*, teachers too often find themselves as condemned and abhorred. Comparatively, no one blamed the band for sinking the ship; rather, everyone praised them for making the dire situation better, giving it all they could with virtually no thought for self. Unfortunately, the inverse is too often true for public school teachers in the United States.

Since the release of *A Nation at Risk* in 1983 (U.S. Department of Education [DoE], 1983) and more recently the enactment of No Child Left Behind in 2002, public school

Address correspondence to Audrey Amrein-Beardsley, Mary Lou Fulton Teachers College, Arizona State University, P. O. Box 37100, Phoenix, AZ 85069. E-mail: audrey.beardsley@asu.edu

districts and schools, including their administrators, teachers, and students, have been keeping afloat in unsteady waters, facing another type of peril. They are being held accountable and responsible for nearly everything they do using standardized test scores, in almost complete isolation of other indicators of educational quality, to measure how and to what extent they are advancing pre-kindergarten to Grade 12 (pre-K–12) student learning (Heubert & Hauser, 1999; Nichols & Berliner, 2007; Orfield & Kornhaber, 2001).

The reauthorization of the U.S. Higher Education Act (Higher Education Amendments, 2007) placed even more treacherous icebergs into the dark waters of teacher education. New Title II provisions have been devised to hold U.S. public colleges of education accountable for the teachers they graduate, adding teacher educators to the list of those to be held accountable and responsible for their graduates' impact on pre-K–12 student learning. Again, they are being held accountable and responsible for nearly everything they do and, again, are often using standardized test scores in almost complete isola-tion of other indicators of program quality but this time using graduates' teacher licensure test scores and the standardized test scores of their graduates' pre-K–12 public school students. In other words, public colleges of education must either evidence that they are preparing teachers who can teach, or more descriptively pass teacher content tests and help their students pass standardized tests, or face the consequences (Cochran-Smith, 2001, 2004, 2005; Cochran-Smith & Fries, 2001; Darling-Hammond, 2006a, 2006b; Field, 2008; Good et al., 2006; Grossman, 2008; Hamel & Merz, 2005; Russell & Wineburg, 2007).

The consequences of this have not been fully articulated beyond the ranking of states' public colleges of education (Hamel & Merz, 2005), and none of the consequences antic-ipated have yet been executed. For example, closing colleges of education—one potential and very serious consequence—seems coun-terintuitive; yet, the threat to close them remains.

In response, faculty and other personnel from colleges and universities in several states have banded together, holding tight to the rail, and have begun efforts to prove their worth and save themselves. These con-sortiums include Arizona, with its Teacher Preparation Research and Evaluation Project (T-PREP); California, with its California State Universities' Center for Teacher Quality (CTQ) and its Performance Assessment for California Teachers (PACT) projects; Louisiana, with its Teacher Preparation Accountability System; Ohio, with its Teacher Quality Partnership (TQP); South Carolina's project as conceived by its legislature (H. 3055, 2007); Texas, with its Center for Research, Evaluation, and Advancement of Teacher Education (CREATE); Virginia, with its Virginia Improves Teaching and Learning (VITAL) program; and Wisconsin, with its Grassroots Teacher Quality Assessment (TQA) model. Individual institutions have also begun these efforts, including Western Oregon University, the University of California Los Angeles (UCLA), the University of Chicago, and others, several of which have part-nered with private entities including Teachers for a New Era (TNE), a project developed by the Carnegie Corporation of New York, the Annenberg Foundation, and the Ford Foundation. Multi-program consortiums have also united, including The Renaissance Group, Teacher Policy Research, and The American Association of Colleges of Teacher Education (AACTE) Standards-Based Teacher Education Project (STEP).

Assembled together are educational researchers, education leaders, teacher edu-cators, and other stakeholders who are developing assessment models by which the effectiveness of their teacher education programs can be meaningfully assessed, all of whom aim to improve the way teachers are prepared and to ultimately advance student learning (Berry, Fuller, & Reeves, 2007; Center for Teacher Quality, 2007; Cochran-Smith, 2001, 2005; Darling-Hammond, 2006a; Hamel & Merz, 2005; Noell, 2006; Rubenstein, 2007; Russell & Wineburg, 2007; Wineburg, 2006; R. J. Yinger, Daniel, & Lawton, 2007). All involved are illus-trating devotion and loyalty to each other and the teachers who have "boarded their

ships," hoping not to perish without the rigorous scientific research and evidence (Cochran-Smith, 2004) they must have to prove their worth and save their institutions and themselves. "Courage and compassion joined," they too are making the process of teacher education evaluation complete in very collaborative and remarkable ways (Scarth, 2009, p. 199).

In short, the U.S. public education system has loads of passengers (approximately 50 million students enrolled in public schools) who need to be secured in quality lifeboats (approximately 15,000 school districts)—boats guided by able-bodied personnel (approximately 3 million public school teachers). As the world and social situations continue to thrust even more treacherous obstacles in the way of these boats, there is a contingent of band leaders and members attempting to support, improve the quality of, and aide the attempts of teachers. This group is the aforementioned education community, charged with providing sound judgment in times of need and not simply giving up and telling passengers that they should've booked their fares on the *Carpathia* or *Californian* (nearby ships that rescued *Titanic's* survivors).

In this article, the efforts of one of these consortiums, T-PREP, are presented. Described are the journey, the distance covered thus far, and the icebergs, so to speak, hit that have ultimately caused a need to radio for help, proceed more cautiously, but still make progress with more and more passengers boarding and more shared understandings being built.

The Setting

The state at the focus of this study is the sixth largest state in the United States in terms of area, the 18th largest in terms of total population, and the 8th most urban state. Its capital city is the fifth largest in population, geographically exceeding Los Angeles, and has just over 1.5 million residents (City of Phoenix, 2011). The state's population has grown 16% compared to the national average of 5%, and public school enrollment is up 18%

since 1999 compared to the national average of 4% (Discovery Phoenix). Within state borders there are approximately 220 school districts; 2,500 elementary, middle, secondary, and charter schools; 71,000 full-time teachers; and over 1,000,000 students (National Center for Education Statistics [NCES], 2007).

The majority of its citizens believe that the state is among the worst in the nation in terms of education and the welfare of its schoolchildren (Arizona Education Association, 2008). According to *Education Week* (EPE Research Center, 2008) reports, the state ranks 49th of the 50 states on equity and spending indicators, 46th for "chance of success" indicators, and 40th for efforts to improve teaching. According to the Annie E. Casey Foundation (2008), the state ranked 46th of the 50 states on the percentage of children living with a head of the household who was a high school dropout, 45th on the percentage of young adults enrolled in higher education, and 45th on the percentage of children living with immigrant families. In 2007 the state ranked 47th for the percentage of children who were English-language learners (ELLs), 47th for the percentage of children living in poverty, and 46th on the proportion of teenagers who were high school dropouts. In addition, according to the NCES (2007), the state's public school students consistently rank among the worst in the nation (bottom quintile) across grade levels and subject areas on the National Assessment of Educational Progress (NAEP; NCES, 2009).

The Vessels

Within state borders, there are three public universities, all with college of education campuses at central and satellite sites across the state. Collectively, they graduate approximately 3,000 teachers each year. These are the colleges involved in T-PREP.

What project participants are learning, because of the breadth of this effort, should help to inform others as they embark on similar journeys. What the field needs right now are studies that include large public institutions (like the Ohio Teacher

Quality Partnership and the California State Universities' initiatives) and possibly large private institutions (e.g., University of Phoenix) to inform others about their experiences in evaluating, redefining, and improving the quality of teacher education in the United States and beyond. Such efforts will greatly inform others as to how they might avoid certain perils as they too embark on similar journeys (Grossman, 2008). The remainder of this article presents the decisions made, challenges faced, and results of one part of this large-scale, statewide effort.

The Maiden Voyage

The initial 2 years of T-PREP have been challenging as well as rewarding and have resulted in seven significant accomplishments: First, project leaders brought together the deans and select faculty from each of the three universities' colleges of education to participate in this project. Because these deans also lead each of the university's satellite programs, their involvement facilitated the inclusion of all of the public teacher education programs in the state. Although college leaders and faculty had collaborated before, this was the first time they came together to collectively and comprehensively embark on such a colossal and high-profile research project. This was not a simple feat but essential to the effectiveness and sustenance of this important research.

Second, project leaders involved other major members of the "band" to help inform this process; that is, leaders from the state department of education, the governor's office, the local teachers' union, the local school administrators' association, leading school administrators and practitioners, and the like. Project participants believed that it was most important to include leaders from each educational constituency throughout the state, particularly to come to shared understandings about everything from conceptualizing what it means to be an effective teacher to coming to consensus about the methods needed to measure teacher education quality. Again,

this was not a simple feat but essential, particularly so that all constituencies could begin building a culture of evidence and illustrate to the public that, together, the state's public colleges of education were prepared to use data in summative and formative ways and to inform decisions, transformations, advancements, and enhancements

Third, project leaders and participants collectively created a conceptual framework— seven "beyond excuses" imperatives—to help conceptualize how this research should and would be conducted; a conceptual framework that should have a significant impact as more public colleges of education at many levels are required to advance research in this area.

Fourth, project leaders and participants collectively began to plan a model to research and evaluate teacher education program quality, particularly in terms of graduates' impact on pre-K–12 student learning, the main and ultimate purpose for this research. Project leaders created a concept map for evaluating the colleges of education and addressed the details of data collection, analysis, and dissemination, all of which serve to advance this project.

Project leaders agreed that all research endeavors and instruments would be aligned with the National Board for Professional Teaching Standards' (NBPTS) five core propositions.[1]

Fifth, at one of the universities involved, the largest of the three participating colleges combined all colleges and divisions of education into one, which allowed for an economy of scale in using resources, providing standardized and consistent coursework and experiences, and reaching all students and potential employers with a unified message. Additionally, across the entire state, project participants have begun following through with this trend to standardize common syllabi, vertically and horizontally equate courses and course components (e.g., via student signature assignments), standardize instruments (e.g., exit and entry surveys, field experience instruments, etc.), and appoint course coordinators to ensure quality across similar courses.

Along with this came the obvious increase in opportunities for professional development.

Sixth, project leaders worked with state department personnel to develop and pilot electronic institutional recommendations (IRs). In Arizona graduating teachers all need an IR to become state-certified teachers. Putting this online will immensely benefit the project because the data collected are needed as a part of the state's longitudinal database and the data will permit all involved to follow teachers from their teacher education programs into the field and beyond. As part of building this database, project leaders also designed a Milestones Project in which all information related to teacher education students is housed in a central online source.

Seventh, project leaders and participants continue to recognize the changing needs and demographics of the state (previously described) and the urgency for the state's colleges of education to ensure that its graduating teachers feel they have the tools to enter the school system and effectively promote pre-K–12 student success. This project has provided a wealth of information to improve the day-to-day operations and formative decisions within the colleges of education.

Initial Results

This information stems from the growing amount of insight solicited from student exit surveys—nearly 1,800 of the state's graduating preservice teachers have been surveyed since Spring 2008.[2] And with this ever-increasing sample, some consistent results are emerging that highlight some of the strengths and weaknesses in terms of the teacher education programs involved, including:

1. Overall, students consistently report being well prepared for the teaching field. Students report that their programs prepare them well for the teaching field; they believe their faculty are experts in the field of education and are committed to teacher preparation; and they believe their

programs are of high quality. Responses also indicate that programs are preparing students in connection to the five core propositions of the NBPTS.
2. Almost all (98%) students report being confident in their ability to use technology in the classroom. However, the program with the most respondents received numerous technology awards, which surely biases this number.
3. Over 90% of all respondents indicate that they anticipate teaching for more than 5 years, although we know that teacher graduates often overestimate how long they will actually be teaching in the field (Berry, 2004; Ingersoll, 2002; Ingersoll & Smith, 2003; Wayne, 2000).
4. Respondents express strong desires regarding their future employment, including: good benefits over a high salary; working in a school near their home; and working in a public school (for similar discussions, see Glennie, Coble, & Allen, 2004; Gritz & Theobold, 1996; Hanushek & Rivkin, 2003; Ingersoll, 2002; Murnane & Olsen, 1989; Tye & O'Brien, 2002).
5. Students also provided valuable information to further improve their education. Specifically, students request a stronger connection between their coursework and internships—connecting their training to practice (see also Dean, Lauer, & Urquhart, 2005; Wilson, Floden, & Ferrini-Mundy, 2002)—consistency among information providers (e.g., instructors, advisors, field experience); and more subject area expertise.

The participating teacher education programs are already responding to these requests and using this information by aligning programs internally; aligning field experience placements; streamlining coursework, advising, and field experience requirements; aligning reporting mechanisms; and considering how colleges of arts and sciences may provide more and reformed content courses for preservice teachers.

As stated, numerous lessons have been learned from this project that may be generalized. But global lessons have also been

learned about conducting these types of large-scale studies.

Safety Check: Locate the Lifeboats

Lifeboat 1: Align the Instruments With Nationally Accepted Standards

In terms of the lessons that project leaders and members learned from the exit survey, they discovered that they floated a good, reliable, and valid survey (Barnett, Amrein-Beardsley, & Koerner, 2009), one that stands apart from the others in that it is the first to be aligned with the NBPTS's five core propositions. There were some questions that had to be changed after the initial pilot (e.g., those that raised flags in the reliability analysis that were written in reverse prose, inversely coded, and likely confused respondents), but overall the instrument fared well and arguably captured the construct of the teacher graduate's perceived effectiveness of the program.

Lifeboat 2: Avoid Fatigue and Duplication

If such survey instruments are administered while other faculty and staff are administering other research/evaluation survey instruments, response rates across surveys administered simultaneously decrease due to a survey surplus, which creates confusion and burnout. Teacher education programs conducting this research might either prohibit the administration of multiple surveys during the time that graduate surveys are administered, and/or graduate surveys might be institutionalized. This more dramatic approach would require all graduates to participate in these exit surveys in order to receive, for example, their IRs (as was done in this project), their final grades, their transcripts, and/or their diplomas.

This point is not necessarily unique at first glance; that is, intuitively a college with a single instrument was able to produce the highest response rate. Perhaps the most telling aspect

about this point and the reason to state it is that many faculty and administrators are unaware of how many surveys, papers, and final documents graduating teachers must produce. The lesson learned here is to locate all of these documents and attempt to streamline the process for students, reducing duplication and the time and effort required of them when graduating.

Lifeboat 3: Keep It Simple

Although most researchers value free-response data and feedback that is not forced into reductionistic categories, open-ended data cause drag. This occurs because the process is not as straightforward as it might seem, largely because of the size and diversity of the responding population. As researchers code data and, more specifically, attempt to code responses by teacher education program, then by division/department, by major, and by different combinations of respondents (e.g., students majoring in bilingual education have distinctly different recommendations for program improvement than elementary education majors, although they all exist within the same division), this drowns the results, not to mention the researchers keeping their heads above the water analyzing these voluminous data. Although partitioning these data is not impossible, it is more burdensome than expected. The ways in which data are to be disaggregated, analyzed, and then disseminated to decision makers is an important conversation to have, especially before researchers simply add numerous free-response items because they provide richer data and before researchers falsely presume that they can analyze these data mostly at aggregate levels.

Lifeboat 4: Stop, Collaborate, and Listen

Those involved in this project discovered that collaboration and trust are key, especially when determining how and what data should be collected, how data should be analyzed, and how results should be disseminated. Those involved need to be willing to listen

and accept constructive criticism, and when program flaws or weaknesses are revealed, program personnel and administrators should be accountable for making the needed changes. For years, teacher educators have known that students want teacher education coursework that is more applicable, in real time, to the field (Boyd et al., 2007; Dean et al., 2005; Wilson et al., 2002). So why is it that this continuously emerges as a student complaint? It is likely that this problem persists because program personnel and administrators are not being held accountable for using these data in pragmatic, transformative, and evolutionary ways.

Though this lesson is not unique, it is worth reiterating. In 2000 the U.S. Department of Education established the National Awards Program for Effective Teacher Preparation. It was designed to award programs proven effective on six criteria, the last of which was to honor researchers who established a culture of evidence, one that supported the use of data for the evaluation of their teacher education programs (Institution of Education Sciences [IES], 2003; see also Wineburg, 2006). The programs awarded were honored because personnel defied what is too often the case—teacher educators who do not work together to conduct this research. The research community, in particular the educational research community, needs to avoid isolative efforts, which ultimately confuse and disenfranchise schools and teachers and, most important, may not improve upon the student experience or levels of student achievement (Wineburg). "Success seems to hinge on a profession's ability to clearly delineate a focus for its work" (R. Yinger, 1999, p. 108).

Though in this project stakeholders welcomed the opportunity to professionalize the field across the state and worked toward this goal collaboratively with project leaders with whom they built trust and a transparent, research plan focused on student learning, project members still faced some complications. But who said that social science research, especially large-scale educational research, was easy (Berliner, 2002)? This reiterated to project leaders and participants

that collaboration was even more essential than they conceived and more time consuming. This was important if they were to ensure that all involved had the opportunity to come to shared understandings and build consensus as they proceeded, collectively and cohesively.

Lifeboat 5: Determine a Collective Course

In New York City, the school system embarked on such a journey, collaboratively, to evaluate the 75 different teacher education programs and 20 distinct teacher education institutions where the majority of their teachers are prepared. Project researchers noted that "a number of disparate institutions, each with different and sometimes competing interests, had to agree to support this project, including agreeing to share data" (Boyd et al., 2006, p. 157).

As discussed, collaboration is essential, but it is also essential to share data, an imperative that needs to be discussed and deliberated at the outset of such research and evaluation efforts. This is particularly important if participating colleges of education are public entities. Negotiating how evaluative data are collected and shared is no trivial task.

Negotiations are also necessary when those involved in this research think about how the data collected might be analyzed. Analyzing what might seem as simple descriptive data (e.g., responses to an exit survey) is much more complex than what collaborators might expect, especially given the major and minor differences within and between the populations of students who are being examined. What should not be overlooked is that populations of students who oftentimes seem like they might be very similar—they all want and have similar desires to become teachers, for example—are distinctly different and almost certainly vary systematically across programs. When evaluated, the between- and across-group differences need to be controlled for and contextualized before simple, causal-comparative analyses can yield valid results (see also Boyd et al., 2007). Because of the

complexities inherent in educational research (Berliner, 1976; Boyd et al., 2006; Russell & Wineburg, 2007; Starkman, Bellis, & Olsen, 1979; Wenglinsky, 2002; Wineburg, 2006), those involved must discuss what analytic methods will yield the most valid and authentic and least misleading results before analyses begin.

This implies adherence to standard practices in the field including reporting instrument reliability and validity and inferential statistics including sample sizes, means, standard deviations, p values, effect sizes, and confidence intervals (when possible; Zientek, Capraro, & Capraro, 2008). But this also includes discussions about whether data should be interpreted in an isolative manner or comparative analyses are warranted, especially if traditional and alternative programs are involved in such collaborations.

According to Boyd et al. (2007), a cardinal rule when conducting such large-scale evaluations is that "studies should compare practices across institutions to identify effective practice" (p. 3; see also Wineburg, 2006). To take it one step further, Hess (2001) promoted this practice because such comparisons will allow public colleges of education to be exposed to "the cleansing waters of competition" (p. 22; see also Finn, 2001; Levine, 2006). However, with comparisons come market forces, making programs susceptible to competition, ranking systems, and the promotion and pursuit of self-interests (Stone, 1997; see also Grossman, 2008). Practices requiring collaboration for the greater good, driven by democracy, and for the promotion and pursuit of student interests are unfortunately seen as the exception (Cochran-Smith, 2005; see also Cochran-Smith, 2001). Striking the balance between appropriate comparisons is one of the issues that will test the collaboration and trust, which is why the time invested in this area will pay dividends throughout the project. It is necessary that these conversations help to yield synergistic results that, with integrity and transparency, provide best evidence as to what it means to be prepared by a high-quality teacher education program.

Lifeboat 6: Constructive Criticism

In 1976, David Berliner noted that those who evaluate teacher education programs too often suffer from "ostrichism" (p. 5), a disease afflicting educators who, when study results are unexpected or blemishes exposed, stick their heads in the sand, hoping that problems will pass. Teacher educators teach students to be reflectivepractitioners and likewise should have no issues with being thoughtful and critical of their own teacher education programs. Though it is true that many flaws can be explained away—for example, when distinctly different samples of students respond in significantly dissimilar ways about program quality—it is also true that nothing will change unless after imperfections are revealed and understood in context they do not inform change. Though "teaching is, after all, a very complex set of events which cannot be easily understood" (Berliner, 1976, p. 12) and educational research is the hardest science of all (Berliner, 2002), teacher educators who research their programs need to remember that in order for programs to improve, criticality in practice and theory is imperative (see also Cochran-Smith, 2005). Otherwise, the purpose for evaluating teacher education programs is sunk.

Thus far, one college in one of the three universities involved has rewritten their mission, as largely informed by the results of this project. The college now intends to prepare measurably effective teachers that positively impact pre-K–12 student achievement and growth, collaborate with school districts to implement meaningful pre-K–12 reform, challenge historical norms, and lead reform in teacher education. It is changing to a 3 years of content plus one year of pedagogy model and is involving other colleges (e.g., arts and sciences, humanities, engineering) to strengthen course content. It is also downsizing, raising selection and admission standards, offering many more professional development opportunities, enhancing program evaluation initiatives, and acculturating others with the reformed mission and objectives of the college.

Lifeboat 7: Follow Up

Former president of Teachers' College, Columbia University, Arthur Levine (2006) stated that "too often teacher education programs cling to an outdated, historically flawed vision of teacher education" (p. 1) removed far from practice. Haberman (2004) claimed that nothing ever found from research about colleges of education has been applied to practice. Further, Haberman posited that teacher education programs are not based on data but reflect custom, tradition, and the convenience of faculty.

The final lesson learned is that after study results call for change, teacher educators should actually make the required adjustments (see, for example, Grossman, 2008). Rarely do those who evaluate teacher education programs follow up with those in charge to determine whether changes were made and, more important, if not, why not? Though it is true that federal and state departments of education and university accreditation units like to see that teacher educators are evaluating the quality of their programs, these evaluative efforts become merely symbolic if left at that, when results are not used to inform instrumental change. Whether these evaluative efforts are merely episodic or transformative must be determined early in the process if successful efforts are to be worth the extraordinary time and effort spent.

In addition, follow-up is vital when multiple stakeholders are involved. Without the ability to be forthright regarding how results are used, particularly across collaborative units, evaluative endeavors become less transparent, layers of information are lost, and programs simply deepen their roots averting further change; that is, increased ostrichism. Participation or compliance will do nothing for either small- or large-scale efforts unless data are used to inform decisions and program excellence (see also Berry et al., 2007). There has to be some sense of accountability within and across collaborators to ensure change occurs.

Navigating New Waters

As the maiden voyages have reached their ports, the project team continues to prepare for subsequent crossings. Project leaders recently completed an assessment map that will help them traverse future research and evaluation journeys, all of which should ultimately help them answer their main research question: To what extent are their graduates effective as teachers? Eventually all data collected will be analyzed longitudinally, year after year, to better investigate students' initial expectations, what they feel they have learned in the end, and how they value what they learned during their teacher education program, once practitioners in the field. (For more information please see the T-PREP Assessment Model at: http://education.asu.edu/projects/t-prep/assessment-model).

One of the next steps is to also work with the private teacher education programs to come to even more shared understandings about what works to promote pre-K–12 student learning and how teachers in the state can be best prepared. This is precisely why including key educational leaders from the state department to the teacher's union is crucial, given the massive complexity of this mission and how many lives should be affected by this work.

Beyond state borders, project researchers and participants substantiated that what is needed is a national framework that will help them and others navigate these waters (Russell & Wineburg, 2007; see also Wineburg, 2006). While those involved in T-PREP are learning to plot their course as they drive the boat, surely others are too. It would be much easier for all if there existed a series of maps drawn from such expeditions, so to speak. This would not compromise local evaluative efforts or wipe out local control but would help others steer through similar waters in more educated ways; that is, help others avoid the obstructions that have rocked the boats of those pioneering these journeys. This piece is one such map, for one such journey. Colleges of education across the United States

truly are at the zero hour. The time for simply ignoring issues or making excuses has passed.

Notes

1. Proposition 1: Teachers are committed to students and their learning.
   Proposition 2: Teachers know the subjects they teach and how to teach those subjects to students.
   Proposition 3: Teachers are responsible for managing and monitoring student learning.
   Proposition 4: Teachers think systematically about their practice and learn from experience.
   Proposition 5: Teachers are members of learning communities. For more information, see: http://www.nbpts.org/the_standards/the_five_core_prop ositio
2. For more information about these studies, including methodological information, the survey instrument, and the data pertaining to its validity and levels of reliability, see http://education.asu.edu/projects/t-prep ▣

References

Annie E. Casey Foundation. (2008). Kids Count Data Center. Retrieved from http://www.aecf.org/

Arizona Education Association. (2008). Comparative rankings. Retrieved from http://www.arizonaea.org/profiles.php?page=32

Barnett, J. H., Amrein-Beardsley, A., & Koerner, M. (2009). The state of teacher preparation programs in Arizona: Year two results from the Teacher Preparation Research and Evaluation Project (T-PREP). Retrieved from http://tprep.asu.edu/T-PREP_Year_2_Report_web.pdf

Berliner, D. (1976). Impediments to the study of teacher effectiveness. Journal of Teacher Education, 27, 5–13.

Berliner, D. (2002). Educational research: The hardest science of all. Educational Researcher, 31(8), 18–20.

Berry, B. (2004). Recruiting and retaining "high quality teachers" for hard to staff schools. NASSP Bulletin, 88(638), 5–27.

Berry, B., Fuller, E., & Reeves, C. (2007). Linking teacher and student data to improve teacher and teaching quality. Retrieved from http://www.dataqualitycampaign.org

Boyd, D., Grossman, P., Lankford, H., Loeb, S., Michelli, N., & Wyckoff, J. (2006). Complex by design: Investigating pathways into teaching in New York City schools. Journal of Teacher Education, 57(2), 155–166.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2007). Examining teacher preparation: Does the pathway make a difference? Retrieved from http://www.teacherpolicyresearch.org/portals/1/pdfs/Examining_-Teacher_Preparation_Exec_Summary.pdf

Center for Teacher Quality. (2007). Teacher preparation program evaluation based on K–12 student learning and performance assessments by school principals. Retrieved from http://www.calstate.edu/teacherquality/documents/teacher prep_eval_results_principals_assessment.pdf

City of Phoenix. (2011). Community trends and profile. Phoenix, AZ: City of Phoenix. Retrieved from http://phoenix.gov/CITYGOV/stats.html

Cochran-Smith, M. (2001). The outcomes question in teacher education. Teaching and Teacher Education, 17(5), 527–546.

Cochran-Smith, M. (2004). Teacher education in dangerous times. Journal of Teacher Education, 55(1), 3–7.

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

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

Darling-Hammond, L. (2006a). Assessing teacher education: The usefulness of multiple measures for assessing program outcomes. Journal of Teacher Education, 57(2), 120–138.

Darling-Hammond, L. (2006b). Constructing 21st-century teacher education. Journal of Teacher Education, 57(3), 300–314.

Dean, C., Lauer, P., & Urquhart, V. (2005). Outstanding teacher education programs: What do they have that the others don't? Phi Delta Kappan, 87(4), 284–289.

EPE Research Center. (2008). Arizona: Quality Counts 2008 (pp. 1–14) Retrieved from http://www.edweek.org/ew/toc/2008/01/10/index.html

Field, K. (2008, September 5). Congress shows colleges they're not off the hook on accountability. *The Chronicle of Higher Education.* Retrieved from http://chronicle.com/weekly/v55/i02/02a03203.htm

Finn, C. E. (2001, July 30). How to get better teachers—And treat them right. *Hoover Digest, 3.* Retrieved from http://www.hoover.org/publications/digest/3459411.html

Glennie, E., Coble, C. R., & Allen, M. (2004). *Teacher perceptions of the work environment in hard-to-staff schools.* Denver, CO: Education Commission of the States. Retrieved from http://www.ecs.org/clearinghouse/55/87/5587.doc

Good, T. L., McCaslin, M., Tsang, H. Y., Zhang, J., Wiley, C. R. H., Rabidue Bozack, A., & Hester, W. (2006). How well do 1st-year teachers teach: Does type of preparation make a difference? *Journal of Teacher Education, 57*(4), 410–430.

Gritz, R. M., & Theobold, N. D. (1996). The effects of school district spending priorities on length of stay in teaching. *Journal of Human Resources, 31*(3), 477–512.

Grossman, P. (2008). Responding to our critics: From crisis to opportunity in research on teacher education. *Journal of Teacher Education, 59*(1), 10–23.

H. Bill 3055. 117th Cong. (2007). Retrieved from http://www.scstatehouse.net/sess117_2007-2008/bills/3055.htm

Haberman, M. (2004, February). *Alternative certification: Intended and unintended consequences.* Paper presented at the first National Center on Alternative Certification Conference, San Antonio, TX.

Hamel, F. L., & Merz, C. (2005). Reframing accountability: A preservice program wrestles with mandated reform. *Journal of Teacher Education, 56*(2), 157–167.

Hanushek, E. A., & Rivkin, S. G. (2003). *The economics of school choice.* Chicago, IL: University of Chicago Press.

Hess, F. (2001). *Tear down this wall: The case for a radical overhaul of teacher education.* Washington, DC: Progressive Policy Institute.

Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation.* Washington, DC: National Academy Press.

Higher Education Amendments of 2007. (2007). In GovTrack.us (database of federal legislation). Retrieved from http://www.govtrack.us/congress/bill.xpd?bill=s110-1642

Ingersoll, R. M. (2002). *Why do high-poverty schools have difficulty staffing their classrooms with qualified teachers?* Washington, DC: Center for American Progress and the Institute for America's Future.

Ingersoll, R. M., & Smith, T. (2003). The wrong solution to the teacher shortage. *Educational Leadership, 60,* 30–33.

Institution of Education Sciences. (2003). *Systematic evaluation for continuous improvement of teacher preparation.* Retrieved from http://www.mcrel.org/PDF/TeacherPrepRetention/5041RR_SystematicEvalTchrPrep-Intro.pdf

Levine, A. (2006). *Educating school teachers.* Retrieved from http://www.edschools.org/pdf/Educating_Teachers_Exec_Summ.pdf

Murnane, R. J., & Olsen, R. (1989). The effects of salaries and opportunity costs on length of stay in teaching: Evidence from Michigan. *Review of Economics and Statistics, 71*(2), 347–352.

National Center for Education Statistics. (2007). *State profiles: The nation's report card.* Retrieved from http://nces.ed.gov/nationsreportcard/states/

National Center for Education Statistics. (2009). *State Profiles: Arizona.* Retrieved from http://nces.ed.gov/nationsreportcard/states

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools.* Cambridge, MA: Harvard Education Press.

No Child Left Behind Act of 2001, Pub. L. No. 107–110, 115 Stat. 1425. (2002). Retrieved from http://www.ed.gov/legislation/ESEA02/

Noell, G. H. (2006). *Assessing teacher preparation program effectiveness: A pilot examination of value added approaches.* Retrieved from http://asa.regents.state.la.us/TE/digest.pdf

Orfield, G., & Kornhaber, M. L. (Eds.). (2001). *Raising standards or raising barriers? Inequality and high-stakes testing in public education.* New York, NY: The Century Foundation Press.

Rubenstein, G. (2007). Building a better teacher: Confronting the crisis in teacher training. Retrieved from http://www.edutopia.org/building-a-better-teacher

Russell, A., & Wineburg, M. (2007). *Toward a national framework for evidence of effectiveness of teacher education programs.* Retrieved from http://www.aascu.org/pdf/07_perspectives.pdf

Scarth, A. (2009). *Titanic and Liverpool.* Chicago, IL: The University of Chicago Press.

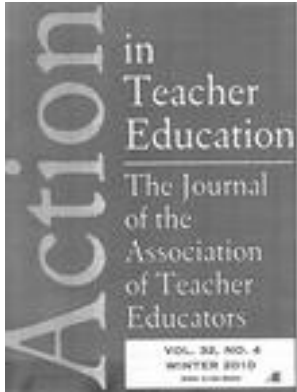Starkman, S. S., Bellis, D. D., & Olsen, C. R. (1979). A heuristic approach to the evaluation

of a teacher preparation program in the light of current realities. *Peabody Journal of Education*, 56(2), 125–132.

Stone, D. (1997). *Policy paradox: The art of political decision-making.* New York, NY: W.W. Norton.

Tye, B. B., & O'Brien, L. (2002). Why are experienced teachers leaving the profession? *Phi Delta Kappan*, 8(1), 24–32.

U.S. Department of Education. (1983). *A nation at risk: The imperative for educational reform.* Retrieved from http://www.ed.gov/pubs/NatAtRisk/index.html

Wayne, A. J. (2000). Teacher supply and demand: Surprises from primary research. *Education Policy Analysis Archives*, 8(47). Retrieved from http://olam.ed.asu.edu/epaa/v8n47.html

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10(12). Retrieved from http://epaa.asu.edu/epaa/v10n12/

Wilson, S., Floden, R., & Ferrini-Mundy, J. (2002). *Teacher preparation research: Current knowledge, gaps, and recommendations.* Retrieved from http://depts.washington.edu/ctpmail/PDFs/TeacherPrep-WFFM-02-2001.pdf

Wineburg, M. S. (2006). Evidence in teacher preparation: Establishing a framework for accountability. *Journal of Teacher Education*, 57(1), 51–64.

Yinger, R. (1999). The role of standards in teaching and teacher education. In G. Griffin (Ed.), *The education of teachers: Ninety-eighth yearbook of the National Society for the Study of Education* (pp. 85–113). Chicago, IL: University of Chicago Press.

Zientek, L. R., Capraro, M. M., & Capraro, R. M. (2008). Reporting practices in quantitative teacher education research: One look at the evidence cited in the AERA panel report. *Educational Researcher*, 37(4), 208–216.

❖ ❖ ❖

Audrey Amrein-Beardsley is Associate Professor in the Mary Lou Fulton Teachers College at Arizona State University. She also serves as co-principal investigator of the Teacher Preparation Research and Evaluation Project (T-PREP).

Joshua H. Barnett is Assistant Professor in the Mary Lou Fulton Teachers College at Arizona State University. He also serves as the director of the Teacher Preparation Research and Evaluation Project (T-PREP).

COPYRIGHT INFORMATION