

# Education Preparation Programs – Tracking Grads

(Alt. Title: *Quality Outcomes ... Who Knows?*)

Gary W. Ritter

Professor of Education Policy and Dean

School of Education

Saint Louis University

Initial Draft: October 31, 2020

Revision Submitted: November 11, 2020

## Education Preparation Programs – Tracking Grads

### 1. Introduction: Why Track Program Completers?

The education policy literature in the 2000s, when education reform was in vogue, was rife with studies focused on teacher quality. Initially, there was a great deal of interest, and debate, around the possibility of performance-based pay, or merit pay, for teachers. Around the same time, education researchers and policymakers engaged in spirited debates over the extent to which evaluations of teacher effectiveness should or could be based on student test scores.

On the heels of these two related policy discussions, observers began to highlight problems in the existing performance evaluations of teachers used in most school districts around the country. The primary problem, according to critics, was that the annual evaluations of teachers were generally not rigorous or meaningful. One of the more prominent reports to point out the flaws of teacher evaluation systems was titled the Widget Effect (Weisberg, et. al. 2009), published by The New Teacher Project. The term “widget effect” describes our practice in K12 education to act as if teacher effectiveness does not vary by individual. That is, by not acknowledging individual differences in classroom effectiveness, we treat teachers as interchangeable parts, or widgets.

This general concern over the weakness of annual evaluations was picked up by the Obama Administration and Secretary of Education Arne Duncan. Using funds stemming from the ARRA (American Recovery and Reinvestment Act), Duncan and his team developed the Race to the Top initiative, or RTT. This program used incentive funding to nudge state actors to develop policies in six key areas, one of which was teacher evaluation. At the time, frustrated by the fact that most systems awarded nearly all teachers with the highest-level rankings, policymakers aimed to reshape evaluation systems so that they might result in more helpful information related to teacher performance. The RTT funds, as a result, were offered to encourage school leaders to attach genuine consequences to more valid and reliable measures of teacher quality.

There are two broad reasons why a school leader engages in the exercise of teacher evaluation. First of all, evaluations (formal and informal) are conducted so that administrators can identify the relative effectiveness of their teachers. School leaders might use the information provided about a teacher’s effectiveness in decisions regarding future employment, professional development, or future leadership positions. A second potential benefit of a thoughtful system of regular evaluation, done well, is that it would act as a sort of professional development and nudge teachers toward more effective practices. Thus, purposeful regular evaluations of teachers could lead to lasting improvements in the human capital of the all teachers in the school.

If the graduates of our Educator Preparation Programs will be evaluated within their schools during their career, we in higher education should be collecting the evaluation data for parallel reasons. First, we should collect such data for accountability purposes in the event that we are asked to document our own effectiveness by entities such as State Departments of Education. Second, we should be collecting these data to nudge our own programs toward internal improvements. The information alone, absent any accountability pressures, may be valuable as otherwise, we may have no real evidence related to our own performance. Indeed, we cannot address problems that we do not see!

If evaluations of individual educators would provide useful information to program leaders and to policymakers, then the evaluation of educator preparation programs would also provide important evidence of program success. The National Center for Teacher Quality certainly shared this sentiment, and thus has regularly reviewed preparation programs in nearly 2,000 institutions in the United States since 2006. The NCTQ reports are based on criteria such as the selection of students and the coursework offered. While these criteria are certainly reasonable, the annual Teacher Prep Reviews have also been widely criticized for a variety of things, including the data gathering strategies (collection and content analysis of syllabi) and accompanying inaccuracies<sup>1</sup>. Moreover, studies estimating the quality of educator preparation programs based on the so-called value-added measures of teacher effectiveness of their graduates did not align well with NCTQ ratings of program quality<sup>2</sup>.

All of this background goes to show that, while there is clearly no consensus on how we should rate programs or schools or teachers, education stakeholders have consistently shown interest in trying to figure out how best to evaluate the performance of educators and to assess the effectiveness of teacher prep programs. As stated above, there are two reasons to engage in the practice of meaningful evaluation of the educator preparation programs: external accountability and internal improvement.

As a sitting Dean of the School of Education at Saint Louis University, I am most interested in evaluation for the second reason. In my view, the best indicator of the effectiveness, or lack thereof, of our School of Education is the work of our graduates when they are in the field. For this reason, the monitoring of program completers beyond graduation day is necessary to understand our effectiveness as a School of Education.

---

<sup>1</sup> <https://nepc.colorado.edu/blog/why-nctq-teacher-prep-ratings-are-nonsense>  
<https://www.edweek.org/ew/articles/2013/07/10/36nctq-2.h32.html>

<sup>2</sup> <https://www.educationnext.org/do-the-nctq-rankings-identify-schools-of-education-that-produce-graduates-who-are-effective-in-the-classroom/>

## 2. Teacher Preparation and Tracking at Saint Louis University

The School of Education within Saint Louis University currently serves just over 500 students in undergraduate and graduate programs. Our undergraduate programs educate just over 110 pre-service teachers in a handful of specializations (including early childhood, elementary, secondary, and special education). Thus, in our undergraduate teacher preparation program, we welcome 20-30 new students each year and graduate roughly the same number.

In addition to our Bachelor's Degree programs, we also educate teachers through our Master of Arts in Teaching (MAT) program. This two-year, alternative certification program serves approximately 40 students (this should grow to 60 next year).

We also educate and certify aspiring educational leaders through our Master of Arts program and our practitioner Doctoral program (Ed.D.) in Educational Leadership. In these programs, we serve approximately 160 students. Finally, the SOE currently serves over 130 doctoral students pursuing research-based Ph.D. degrees. Roughly 50 of these students are pursuing Ph.D. degrees in Higher Education Administration; 80 more are pursuing Ph.D. degrees in various fields such as Special Education, Curriculum & Instruction, and Education Policy.<sup>3</sup>

### A. How Do We Track our Graduates?

Currently, we do not engage in what I would call systematic tracking of completers in any of these programs. However, we do have several informal tracking strategies for our programs training pre-service teachers and for training school leaders. In our pre-service teacher program, we have a "Field Office" that actively works to place teachers in local schools for numerous practicum opportunities during their first three years and then for their student teaching assignments during their senior year. This Field Office is led by a full-time faculty member and a full-time staff member who build strong connections with our students during their time as undergraduates. With this foundation, these two remain in close contact with our graduates in their post-graduation years.

The information we gather on our recent graduates is straightforward and minimal; we keep track of their area of specialization, the institution that employs them, and the specific teaching position (e.g. middle school math) that they occupy in the school. This simple data collection allows us to keep track of the 'employment rates' of the graduates of our programs. Of course,

---

<sup>3</sup> In recent years, we have seen a decline in our number of undergraduates and an increase in our graduate student enrollment.

information such as this is important to students (and to parents of students) considering enrolling in the School of Education at Saint Louis University.

This type of informal survey is feasible given the size of our school and the accompanying small number of annual graduates. For example, using the most recent non-pandemic-affected data, we graduated 24 students with Bachelor of Arts degrees in education in May 2019. Of these 24, we were able to survey 23 of them. Based on our survey, the Field Office reported to me that 16 of our graduates were employed as teachers in the Fall of 2019, 3 graduates were employed in education-related jobs, and the 4 remaining graduates were employed in non-education positions. While this sort of informal data gathering is helpful in that it allows us to assure parents and students that there is almost certain employment awaiting them upon graduation, it does not scratch the surface of the type of data we need to facilitate genuine program improvement.

### B. Barriers to Following our Graduates

If we are not currently doing what we need to do to facilitate program improvement, what is standing in our way?

I will ground my thoughts here based on my experience at my current institution, where I have served as Dean for more than two years, on my nearly two decades as a faculty member in a College of Education in a large state University, and on my observations as an education policy researcher for the past twenty-five years. My hunch is that there are three key barriers; I will start by discussing a barrier that would be cited by most faculty in colleges of education around the country and then move to a couple of barriers that are more nuanced and would be more likely to generate disagreement.

The **first barrier** to gathering meaningful data on the graduates of colleges of education is related to resource constraints. This is certainly the case here at Saint Louis University. Like many other institutions, our School of Education has experienced real declines in undergraduate enrollment in recent years for reasons that are beyond the boundaries of this essay. In any event, the accompanying budget cuts in recent years have been frequent and non-trivial. In higher education, of course, the vast majority of the expenses live in the personnel category, and many of these personnel lines are faculty lines. Faculty lines, whether tenure-track or non-tenure-track, have numerous protections from budget cuts. Thus, when budget cuts fell on our School of Education, we chose a path that many in higher education choose, and first cut non-personnel expenses and then cut staff when we had to reduce personnel expenses. As a result, any dollars that might have otherwise been used to track the success of our graduates – using either outside companies or our own staff – were no longer available.

Moreover, existing faculty were unable to pick up the slack in this area, as their own workloads increased with staff reductions.

However, if we were to genuinely believe this were the only barrier, we would have to be able to point to robust program completer tracking in years when budget surpluses were the norm in higher education. Thus, at some point in the 2000s, there should have been rich and comprehensive initiatives in data gathering on program completers. My sense, from my time at the University of Arkansas and from my read of the education policy literature, is that this did not occur in many institutions. I do not believe that we, as a community of colleges of education, have done a good job of systematically tracking the success of our graduates.

So, if we have not engaged in this type of self-evaluation in times of budget austerity or in times of budget surplus, what is standing in the way? Well, if the problem isn't the resources, perhaps it is the incentive structure.

In my view, this **second barrier** is more internalized in our system and more powerful: this is the barrier related to incentives. That is, there has been no outside push to encourage colleges of education (or other schools within Universities for that matter) to check in on the success and well-being of our graduates post-graduation. Currently, we are incentivized to recruit students and to serve them well while they are students so they come back semester after semester. These incentives appear both in the very direct and obvious form of tuition payments and in the more indirect form of third-party University ratings, such as US News and World Report, which include components related to the "quality" of the entering students and the retention and graduation rates for the students who ultimately enroll.

Conversely, there are not similar incentives for attending to our students' experiences and success post-graduation. Simply, they are no longer paying us tuition and high-profile ranking systems are no longer paying attention to them. The incentive structure might be bit different on other professional schools, such as law and business. For example, high-profile MBA programs clearly advertise the average base salary of their graduates. Thus, MBA programs are at least incentivized to track their program completers through their first post-graduation hire.

Consequently, leaders of education schools face a bit of a conundrum. In times of excess in the past, where we may have had strong enrollment and accompanying resources that would allow for the systematic tracking of program completers, there was no compelling reason to do so. In times of austerity and decreasing enrollments in colleges of education, when we each might have a competitive motivation to gather better data on our graduates to present a compelling case to potential students, we do not have the resources to do so.

Finally, if we were lucky enough to find our way through the maze described above, I am not sure there would be a consensus on what data we *should* be gathering on program completers.

Thus, the **third barrier** I see impeding our path toward a clearer understanding of the success of program completers is a lack of shared agreement on the details of the data we should be collecting.

Indeed, I suspect that the desired data would vary by institution type. Leaders of small and selective colleges of education may not be satisfied with finding out whether their graduates simply found jobs in schools. In my case, at the Saint Louis University School of Education, which is small and relatively selective, we do not have trouble placing students in jobs. Our students, or SLU students, are perceived to be excellent in the classroom and I suspect they normally are. I do not know, however, how much of their excellence is due to their own ability and how much is due to the value we add during their undergraduate years. In a perfect, world, I should be gathering value-added data of one sort or another to ascertain the 'performance' of the School of Education and to facilitate program improvement where appropriate based on the data. It's not clear to me if we should be measuring teacher effectiveness by the growth of their student test scores or through observations of their instructional practices; it's also not clear what we might use as the starting measure of teacher 'ability' as we compute the 'value-added' scores. Nevertheless, this uncertainty about the correct next steps should not lead to inaction.

On the other hand, if I were leading a larger program, perhaps at a non-selective open-enrollment institution, I might be satisfied with simply assessing the extent to which our students successfully landed placements as full-time teachers. Nevertheless, even in this situation, leaders of education preparation programs should be gathering some data on teacher effectiveness above and beyond simply obtaining employment.

Next, I believe that all institutions, regardless of context, should be gathering data on two other important outcomes. Beyond placement of course, we should care about retention over a reasonable number of years post-graduation. Even the simple question of retention can be complicated by numerous circumstances. For example, teachers may enter the workforce, exit for a few years, and then return to the workforce. It is not clear how a retention indicator should consider this situation. In any event, it certainly seems worthwhile to follow our graduates for at least five years post-graduation as one indicator of the effectiveness of educator preparation programs.

Finally, in my view, I would also want to hold Saint Louis University accountable for the number of teachers whom we are placing at schools serving low-income students. I am not sure if this is, or even should be, a goal of all programs. However, as SLU is a Jesuit institution with a social justice mission housed in the center of St. Louis, we bear a special responsibility to work with and for members of our community to support the schools and build the teacher pipeline. I believe that land-grant institutions bear a similar responsibility in serving their states.

To be sure, reasonable people might disagree about the usefulness of (or the need for) collecting many of the indicators I mention above. Some might argue that the students are no longer our responsibility after they leave our campuses. Still others might argue that we have no ability to influence where they work and whether they choose to work with and for low-income students. Lastly, others might find fault with the whole idea of any sort of value-added measures and dismiss that sort of data collection out-of-hand.

This lack of consensus indicates to me that, even if we had the will and the resources to better track our program completers systematically, I am not sure we could agree on the system! Given this potential lack of agreement, there is an opportunity here for an entity such as a State Department of Education to use its authority to step in and create tracking systems that provide helpful information that would serve the dual goals of external accountability and internal improvement.

### C. How Would We Benefit from Better Tracking of our Graduates?

While our field may be a long way from building a consensus around the data we should gather from program completers of colleges of education, there are reasons that we should move in that direction and clear benefits for individual institutions that do improve in his area. From a broader system perspective, particularly in state-funded colleges of education, it would be very helpful to know which institutions are making the most out of the scarce resources allocated to them.

Leaders of state departments of education have an interest in supporting the pipeline of teachers within each state. Currently, I suspect most<sup>4</sup> are flying blind with respect to which of their institutions are graduating future teachers who: (1) take on teaching roles after graduation, (2) take on positions in high-need subject areas or high-need geographic areas, (3) are effective at supporting student achievement, and (4) remain in the field after the first couple of years.

Furthermore, while the teacher workforce is generally a state responsibility, there are national actors and philanthropic actors who have an interest in the four categories described above. These organizations, such as the US Department of Education or major national grantmaking organizations with an interest in teacher quality, would also benefit from knowing more about how well various institutions are doing in each of those areas. Indeed, leaders of institutions would benefit from understanding how well their peers are doing in the various categories; this

---

<sup>4</sup> Some state departments do gather and use helpful data. The earlier citation about NCTQ highlights a study in Florida where EPPs in the state have data on the VAM scores of their graduates. I believe this is the exception.

might help them better identify systemwide needs and develop their own priorities such that they could be addressing gaps in the market.

Of course, prospective students would benefit from this sort of information as this would help inform their decisions in the school selection process. Finally, school leaders, who are the potential employers of these new graduates, would certainly benefit from better knowledge on the relative success of colleges of education in various categories.

### 3. Closing: What Should We Be Doing?

As we think about best practices in this area, or what we should be doing, we should first think of the purposes. Conversations about the purposes of tracking completers should be rooted in the purpose of evaluating educator preparation programs in general. In 2013, the National Academy of Education (NAE)<sup>5</sup> brought together a group of leaders from colleges of education to author a report focused broadly on the topic of the evaluation of teacher preparation programs. As the authors noted in the introduction, there are many such evaluation systems “currently in place, with different purposes and consequences, and a growing need to clarify their advantages and drawbacks” (p. vii). In the report, the authors asked important questions, not unlike the questions asked here. They asked about the strengths and weaknesses of the evaluation systems that existed at the time, interrogated whether the evaluation strategies and the objectives were well aligned, and importantly concluded with a set of guiding principles for the development and use of such systems.

As we think of how we might track program completers, we should think about our reasons for doing so, and the 2013 NAE report helps us here by articulating three fundamental reasons for undertaking such evaluations: (1) providing a basis for accountability where it is warranted, (2) providing needed information to prospective students and potential future employers, and (3) providing information necessary for program improvement. Any suggestions we make should be aligned to one or more of these purposes.

Moreover, as the authors remind us, we must be attentive to the questions of validity. That is, as a Dean of our School of Education, if I am to ask my faculty or staff to engage in data gathering to assess the extent to which we are succeeding in our objective of placing excellent teachers in schools where they are needed, I had better make sure to clearly communicate why I believe the data collected will provide valid measures. Better yet, I should engage faculty and

---

<sup>5</sup> <https://naeducation.org/wp-content/uploads/2016/11/028489-Evaluation-of-Teacher-prep.pdf>

staff in the collaborative process of building the plans for data collection and for converting these data into useful metrics of our “performance” as a college of education.

The NAE authors also remind us to be attentive to questions of relative costs and benefits. As they state on page five, the “limitations of any evaluation system should be weighed against its potential benefits.”

What does this guidance mean to those considering enhancing our work in tracking program completers? I believe this means that we should be humble about the potential usefulness of whatever data we might require and tread lightly as we begin to seek additional work from our staff and additional data from our graduates and from our school partners. We should initially rely on data that are relatively easy to collect, and then test the usefulness of such data. Will we actually make use of this new information? Or will our evaluation report simply sit on a shelf? If the latter is true, then we know that we should not move forward with those initiatives.

Before we send out surveys to all our recent graduates from the past five years, we should first attempt to use state administrative data, which tracks holders of teacher certificates, to ask where our graduates are working, are they being retained, and are they working with high-need students? This last piece of information would be particularly meaningful to me in my role as a Dean at a Jesuit institution.

Before we attempt to ask school leaders to respond to surveys, or to provide observation scores, comparing the effectiveness of our graduates with other teachers, we should first find out if our state provides value-added data at the teacher-level that might help answer key questions.

Before we enter into classrooms and ask students to respond to surveys focused on the level of instruction provided by their teachers, our graduates, we should first directly survey our teachers in the field with questions focused on their own perceived level of readiness for their job. Of course, this sort of work requires a partnership with districts, but it involves a lesser level of intrusion than do observations. This sort of preliminary data might help in facilitating program improvement.

If these less intrusive means of data collection yield meaningful results that are actively used, we might then choose to gather more data through processes that might be more intrusive, but we will need some reason to believe the efforts could bear fruit. In some ways, I am suggesting, as we enter into the important tasks of meaningful data collection and analysis, that we follow our own version of the Hippocratic Oath, imperfectly translated as “First, do no harm!”