

# Carnegie Knowledge Network • What We Know Series on Value-Added Methods and Applications

Webinar 12: What Do Value-Added Measures of Teacher Preparation Programs Tell Us?

## Q&A with Dan Goldhaber

December 11, 2013

**Q: You mentioned the importance of learning more about measures of teacher preparation programs other than value-added. What does research suggest a more well-rounded assessment of teacher training programs might look like?**

A: I would put that into the category of judgment call. Let me take two issues. First, there is a little bit of research, cited in my brief, that looks at the probability that teachers from different training programs stay in the work force. That is certainly what I would consider to be one important outcome that is not value-added. Second, one could also imagine using teacher performance measures that are much more well-rounded, based on, for instance, value-added plus classroom observation, principal assessment, or student perception surveys. These more holistic measures could also be linked back to training programs. The problem with this idea is that in many districts and schools today there is very little variation in those summative measures of teacher performance, so they do not provide useful information about training programs or about teacher quality in general. It may be that that is changing because of Race to the Top and other initiatives, but it is too early to know. Let me add one last thought, there is something new and potentially important called the edTPA (Teacher Performance Assessment). It is designed to assess people as they go through training programs and hopefully guarantee that people graduating from training programs meet a certain threshold of quality. Currently, we do not know much about it. It has been touted, but it is largely empirically untested.

**Q: We have a question about the influence of placement. In your view might the use of value-added in assessing teacher preparation programs also incentive them to develop working relationships only with districts in which graduates are more likely to be effective? What are the possible side effects?**

A: I am not sure that *purposeful* placement is entirely a bad thing. I would distinguish two issues with purposeful placement. One is the issue of specialization. I actually would hope that teacher training program would specialize to some degree to try to meet the needs of local districts since it is pretty clear that student teaching tends to happen near training programs. When student teaching occurs there is a contract between the school district and the training program, so you would really want the student teaching to be centered around the needs of the students the district is serving, and I could imagine that this sort of specialization could be beneficial. More globally, the other thing that could occur would be better matching once teachers enter the labor force (a different type of placement). If we learn that teachers may be differentially effective with different kinds of students, then you would want to put them in jobs where they could make use of their skills. Of course we also might be concerned that value added models are not be adequately accounting for the background of students when assessing teacher training programs. That is a general issue that arises with any kind of value-added model. That issue has not yet been put to bed, but there is now a great deal of research on it so you might want to take a look at some of the other Carnegie Briefs. I will say that the issue of “bias” is less of a problem with assessing TTP than it is with assessing individual teachers. Not that the problem does not exist, but it is less of a problem because you are aggregating over many, many more students.

## Carnegie Knowledge Network • What We Know Series on Value-Added Methods and Applications

### Webinar 12: What Do Value-Added Measures of Teacher Preparation Programs Tell Us?

You get teachers who are teaching students with a variety of backgrounds - advantaged students and disadvantaged, white students and minority students.

**Q: What do we know about the effects of field placements during teaching preparation that precedes the teaching experience?**

A: First, I would urge people to look up Matthew Ronfeldt. He has done some of the best work on this topic. I will say a few cursory things. One is that it looks like doing student teaching in schools that have what Matt characterizes as a high stay-ratio (teachers in a district are likely to stay at a given school in that district) is predictive of how effective they are when they actually end up being responsible for their own classrooms. So, doing student teaching in an environment where teachers want to stay in the school, maybe because of the culture or a wide range of other reasons for a high stay ratio, seems like a good thing. That study is particularly interesting because I would have thought that the more your student teaching experience is similar to your first teaching experience, in terms of the types of students you're teaching, would be predictive. But, at least in Matt's study, that turned out not to be true. Beyond the stay-ratio and the connection between your student teaching experiences and what you are learning, we don't know a lot. We need to learn more about for instance, when student teaching should happen, the curriculum in teaching training, the length of student teaching, the nature of the schools where student teachers are placed, and the mentors student teachers get paired with.

**Q: Have characteristics of these field sites been included in these value-added studies?**

A: They have been included in a couple of studies, but they tend to be based on retrospective assessments of teachers on what they experience. There are a couple of studies now that connect student teaching experiences with teacher effectiveness and include measures about the schools where you did your student teaching experience. Susanna Loeb and some colleagues, including Matthew Ronfeldt, asked in-service teachers about their student teaching experience and about their experiences in teacher training preparation and then connected their answers to survey questions to effectiveness. I have also done a study that connects student teaching experiences that are based on actual administrative data on the student teaching from a number of teacher training institutions. In general I think we are just in the beginning of learning about the features of teacher training.

**Q: Have you been able to make any correlations between teacher effectiveness and ACT and SAT scores?**

A: There is mixed evidence on the direct connection between ACT and SAT scores and student achievement. There is more consistent evidence that teachers who do better on licensure examinations have better student achievement. But it is by no means the case that doing very well on the licensure exam guarantees that you are an effective teacher, nor is it the case that doing poorly on the licensure makes you a bad teacher. There is an association between teacher performance on licensure exams and students achievement. Also, there are a lot of different kinds of licensure tests and the association is not true for all of them.

# Carnegie Knowledge Network • What We Know Series on Value-Added Methods and Applications

## Webinar 12: What Do Value-Added Measures of Teacher Preparation Programs Tell Us?

**Q: From a hiring institution’s perspective, say a district or school, is there any reason they should attend to the difference between selection and training effects of teacher preparation programs?**

A: I think the answer is generally no. I think in general you care about the whole prospective teacher, not just what part of that person is attributable to who they were before they became a teacher candidate. That said, I could imagine in some cases the answer might be yes. For instance, if you knew that you wanted to partner your school district with a teacher training program, then you may well care about the difference between selection and training because the training might be related to what you look for in the partnership.

**Q: The answer might be different at the state level. Are there specific elements that you would recommend states collect to begin to address the issue of selection versus training effects?**

A: Sure, there are lots of things that I would suggest. I think that we could learn a lot more about the distinction between selection and training if we knew information about what was required for applications, what is the selection process among teacher candidates, etc. That information generally resides within universities and is very hard to get hold of. I could imagine a study that would actually distinguish between selection and training effects, if you had really detailed data on the applicants and the selection process.

**Q: To begin examining the difference between training and selection effects, what information about teacher training programs applicants do you feel is most important to have in order to determine the impact of selection on value-added measures of teacher preparation programs?**

A: Some of this would be measurements of teacher cognitive ability or academic proficiency and also their subject matter knowledge, especially for preparation of teachers at the secondary level. I think that the current measures we have are really quite weak. In addition, we do not know much about the interpersonal abilities of people. Some of that can be taught, but also people are quite different in terms of their ability to communicate and connect with others. I would argue, especially at the elementary level, those skills are really important. Part of their job is not just know the subject material, but also to understand what kids do and do not understand, be able to inspire them, and make school exciting and relevant. Those are attributes of people that you may be able to detect through selection processes and could, potentially, be quantified. There are certainly selection instruments that report to do this. I do not know the extent to which selection along any of the aforementioned dimensions occurs when people are admitted into training programs.

**Q: You mentioned that as a teacher progresses they begin to look more like their colleagues. How would you define “looking more like their colleagues”?**

A: The way the model works is that it allows the effect of the training program to diminish over time. Maybe I over stated whether they actually look like their colleagues. Rather, they look more like an individual, instead of a person who graduates from a particular training program. Whether that is about who their colleagues are or not is not embedded in my study. There is a study by Kirabo Jackson that

# Carnegie Knowledge Network • What We Know Series on Value-Added Methods and Applications

## Webinar 12: What Do Value-Added Measures of Teacher Preparation Programs Tell Us?

suggests a proportion, I think roughly a quarter, of the effectiveness of a teacher is about who they have surrounding them as colleagues in their school. So my statement is combining two different studies.

**Q: If we hold teacher preparation program accountable for graduate effectiveness, is there a conflict between the idea that expertise takes times to develop and the fact that program effects decay over time?**

A: The idea that expertise takes time to develop is an empirical question. What I think is interesting about the figure that I presented about the half-life of teacher training program effects is that there was nothing about the model that assumed it would decay. That is an empirically-derived result. One could imagine that a new teacher gets out into the field and relies on a very scripted curriculum. It is only when they gain confidence as a teacher that they remember the lessons they learned in teacher training programs and apply them. If that were the case you could imagine that teacher training programs effects would grow over time as a teacher progressed in her career. It looks empirically (at least based on this one study) like this is not the case. The second thing I would say is that you could still imagine that one gains expertise over time – the many studies showing a return to early career experience certainly suggests this is the norm – but the expertise is not necessarily connected to what you learn while in teacher training.

**Q: At this point, what does the research show has a stronger the predictive effect of value-added: teaching training programs or school and district selection?**

A: That is a great question. I cannot give you a definitive answer to that question because I do not know exactly what the magnitudes are of these two possible effects. I would say that it would certainly depend on the states you look at. The effects in some states, like New York or Washington, look bigger than the effects in, for example, Missouri. So that is one thing to consider. Also, I would say that both of these are areas with tremendously low-hanging fruit when it comes to thinking about workforce improvement. The study that I am thinking about (look for Rockoff et al. in Education Finance and Policy in 2011) looks at selection and suggests that only a small percentage, I think 10 percent, of variation in teacher effectiveness was predictable based on the information that might be collected amongst candidates. I imagine that school districts could do a whole lot better when they are trying to select teachers than they currently do.

**Q: What, in your opinion, is the theory of change in using value-added to assess teacher training programs? How could teacher training programs use the information they receive to improve their program?**

A: If the value added information is a black box, then there is little they can do with it directly. There may, however, be indirect ways that value added influences training programs. For instance, if a program is told that they are less effective than average, they could maybe use it as a launching point for further exploration as to why. But it does not give them direct evidence to say this is what you do to improve. In general I would hope that one of the things that happens with the use of value-added

# Carnegie Knowledge Network • What We Know Series on Value-Added Methods and Applications

## Webinar 12: What Do Value-Added Measures of Teacher Preparation Programs Tell Us?

generally is that people try to learn from each other. Simply putting information out into the public could help to identify exemplars so that there could be more informal learning. For example, in individual value added, a low value-added teacher may go to a high value-added teacher and ask them what they are you doing in their classroom because it seems like it is really working according to this measure. You could imagine that also happening with teacher training programs. But to identify concrete action steps for training programs to get better, we need to get into the features of teacher training programs. For some purposes, like state level accountability, it may be useful to use value-added whether or not it helps teacher training programs get better. I do not of know any evidence that this is the case, but one of the things about value-added that might be useful is that the availability of value-added measures may help to change the types of conversations that go on when other assessments are used. Teacher training programs have to go through an accreditation process. What that looks like now in most places is teams of people from NCATE or the state office going to the program, interviewing professors, looking at requirements, and, based on all this, trying to assess whether it looks like a quality program. Few programs are shut down or found to be failing. That may be about the information generated by the program approval process itself or it may be just the human element: it is hard to give people bad news. I wonder if there is not actually more information out there about which programs are effective or ineffective than currently shows up in these state assessments. If it is the human element, rather than the lack of information about program quality, having value-added allows you to have tougher conversation with a dean, which in turn allows the dean to have tougher conversations with the faculty, because we know that deans do not have all the power in the world and cannot simply change the program at their will.

**Q: Ultimately how will value added itself be evaluated as a means to identify highly effective teachers?**

A: Ultimately, we probably do not care about value-added and student test scores themselves. The only reason I care about these things is because I think that they are good intermediary measures of the things that I really care about, for example, if someone is prepared to be productive down the line. Are schools preparing people to have productive careers and to attend college? Are you more likely to be employed if you have certain teachers rather than others? Is the value added of your teacher connected to your later earnings? Those kinds of outcomes are what we really care about. The way that you are going to validate value-added or any measure of schools, teacher preparation, et cetera, is to see whether interim measures – which we have to use because we do not have 12 or 16 year timelines before we can measure of how an intervention or school is doing – are connected to the long-term outcomes that we really care about.

**Q: Do you know of any similar work that has been done related to principal preparation programs? How might the challenges there be similar or different from those related to teacher preparation?**

A: I don't know of any work like this that has looked at principals preparation programs. I would urge people to look into Susanna Loeb's work because she knows the principal quality work better than I do. That said, my take is that it is harder to judge principals for at least three reasons. Reason one is because

## Carnegie Knowledge Network • What We Know Series on Value-Added Methods and Applications

### Webinar 12: What Do Value-Added Measures of Teacher Preparation Programs Tell Us?

I think it takes time to change schools. Depending on when someone is hired, I do not know to attribute the student achievement from the next year to that principal or not. Secondly, there are sample size issues that make it harder to distinguish principals than teachers, that is, there are fewer principals to compare to one another than teachers who can be compared to one another. Lastly and maybe most important, principals in some cases are pretty constrained. For instance, some principals have more control than others over who is actually on their teaching staff. If you are a principal where you have a lot of control over what is going on in your school, including who is there, then you expect that that principal could have bigger effects, negative or positive, than a principal who is hired and basically stuck with their existing staff and may try to improve what is happening in the school, but is limited to some extent. So when you extrapolate from that, it is more difficult to judge principals than to judge teachers, who can potentially close the door and have students for a set amount of time everyday. All of those reasons make it harder to think about how to evaluate principal training programs.