School Meeting  
March 20, 2013

Present: Frank Adams, Johanna Barnes, Sherry Dorman, Laura Franklin, Joni Irlmeier, Brook Jech, Dennis Lichty, Judy Moeller, Marilyn Mudge, Susan Nordstrom, Jim Ossian, Craig Pease, Tim Sharer, Phyllis Spethman, Bob Sweetland, Sue Sydow, Keith Willis

Absent: Barb Black, Kathleen Conway, Janice Dinsmore, Steve Dinsmore, Terry Hamilton, Daryl Wilcox

Meeting called to order at 8:01 a.m.

1. **Children’s Library Collection Projects**—Charissa Loftis informed the faculty that the library staff are in the process of weeding the juvenile fiction collection. Faculty members have till April 15th to raise any objections. Books that have only been checked out once since 1991 are being pulled. The juvenile non-fiction will be next and faculty will have till the end of the semester to make objections to any of those books pulled. Charissa would like to develop a new system of ongoing maintenance. Due to such a large amount of exam books coming in on a yearly basis, she would like to have the books weeded when they are moved over to the permanent collection (3 years after receipt). If the exam book has not been checked out in those three years, then it will be pulled at that time. This excludes award, honor, series, and easy titles. Charissa will be running a list of pulled books on April 1st and will email this out to all education faculty members to look over. A few faculty members volunteered to go through the list.

2. **Policies & Procedures**—Dr. Lichty explained that Dr. Anderson is trying to compile all policies followed in each school to be sure we are all complying with the policies set forth. The issue is how to start this process. Faculty members are to be thinking about this and what we all need to include in the manual. We will get together after pre-registration advising is complete. It was mentioned that Dr. Blomenkamp had started this process and the faculty met many times to work on this. Maybe we can try to locate what was started and go from there.

3. **Elementary Leadership Conference**—Sheri Filipi has contacted the school to see if we plan to offer the conference in the fall, as done in the past. Right now, the class it was normally associated with is taught by adjunct. If you are interested in taking on this project, contact Dr. Lichty. This is a large project and many faculty members don’t have time to get this done, and asking students to take on one more project is difficult.

4. **Search Committee Updates**
   a. **Dean**—Dr. Willis stated that we are bringing a candidate (Dr. Dave Braun y Harycki) to campus this upcoming Monday and Tuesday. Faculty will have an opportunity to meet him at 2:30 p.m. in room 321
   b. **SPD**—Casey Hurner was hired and is already working with Dr. Franklin on fall courses.
   c. **CSL**—Two candidate have come to campus and recommendations will go to the VPAA this week. Dr. Willis informed the faculty that they are getting an additional faculty line for next year, so they may be able to hire both candidates.
   d. **EDU**—Craig stated that they have had 50 applicants for this position and they have narrowed it down to 6 candidates for phone interviews. These phone interviews will take place this Wednesday, Thursday, and Friday.
5. **Graduate Studies**—Dr. Sharer informed the faculty that we have had a positive response to the Reading Specialist mailing that was sent out. We have 6 people that have declared this program so far. Also mentioned, are the NENTA informational meetings today and tomorrow from 4:00-5:00 p.m. in room 218. Dr. Sharer asked that the faculty inform any students that have completed 60+ hours, have a 3.3 GPA, and are eligible to take Level 2 courses, about NENTA and the informational meetings.

6. **CLSP Update**—Dr. Wilcox is recovering from a broken rib and is doing well. She hopes to return to campus by the end of the week. Dr. Conway will be on sabbatical in the fall, so Dr. Willis is trying to get all the counseling courses covered. The department is starting to complete a follow up survey for alumni from the last 3 years.

7. **Field Experience Update**—All the pre-student teaching placements were completed last week. This was a very large task (350+ placements). There were 92 student teachers this term; however 2 have been removed from student teaching. One will be offered a second opportunity, and the other one was just recently removed, so nothing has been decided yet.

8. **Value Added**—Craig Pease handed an article out by Douglas Harris (see attachment). This is who Craig heard present and has gotten a lot of his knowledge about “value added” measurement from. The NCLB schools report test scores based upon attainment (what achievement they reach at a point in time). NDE has started to look at growth in their model (amount of academic gain or progress a student makes over a period of time). By using statistical data and analysis, we can compare where we predict a certain group of students will perform with how they actually perform. Many states are building this “value added” system into their teacher evaluations. Harris believes that this system is not stable enough to evaluate teachers partially because we are not statistically advanced enough to make this work. As with any measurement system, there are many factors that come into play.

9. **Upcoming Dates**—Dr. Lichty reviewed the following dates with faculty:
   a. March 23—Just Juniors (Dr. Adams and Josh York)
   b. March 25—Pre-Registration Advising Begins (Dr. Lichty will be taking all the campus visits that week), PPC Meeting, TEAC Dinner & Meeting
   c. March 28—Final Sexual Harassment Prevention Training 8:30-10:00 and 1:00-2:30 in GH Auditorium—this is mandatory so if you haven’t attended one yet, you must attend one of these two.
   d. April 1—Spring Break
   e. April 8—Fall Registration Begins
   f. April 10—On Campus Retirement Open House for EDC Retiring Faculty 3:00-5:00 p.m. in Elkhorn Room of the Student Center
   g. April 15th—Off Campus Retirement Party EDC Retiring Faculty 3:30-5:30 p.m. at Miss Mollie’s Coffee Shop, downtown Wayne

10. **Other**
   a. Senator Adams has been selected as our Outstanding Alumni.
   b. Dr. Lichty will be emailing faculty about next summer dates (3 options). Please provide feedback to him and it will be presented at the next Dean’s Council.

The meeting was adjourned, 9:00 a.m.

Minutes Respectfully Submitted by Brook Jech.
Clear Away the Smoke and Mirrors of Value-Added

Accurately measuring performance is important to any organization, including schools. State longitudinal data systems with annual student assessments are expanding the possibilities for measuring school and teacher performance. This can increase the potential for evaluating educators in ways that are fair and improve performance and student outcomes.

Unfortunately, despite almost two decades of the "new accountability" that links student test scores to schools, our performance measurement and accountability systems are still broken. Some of the problems — narrowing the curriculum and teaching to the test — have received widespread attention. Another critical problem is less widely recognized. The "Fundamental Principle of Accountability," as I will call it, is that we should hold people accountable for what they can control. But accountability systems violate this principle when they focus on student attainment, or student achievement at a point in time.

The problem with attainment has to do with a basic fact of human development: Knowledge and skill accumulate over each person's life. What adults know and can do depends on what they experienced as students, which in turn depends on experiences before school, all the way back to birth (Heckman 2006). Kindergarteners have vastly different early childhood experiences and readiness for school. For example, black students start kindergarten with scores that are already 22 percentile points behind white kindergartners (Fryer and Levitt 2004). Schools have not caused these "starting-gate inequalities," because most students haven't set foot in a classroom before kindergarten.

The gaps are so large and persistent that even effective schools don't completely overcome them. Not only are the starting gate inequalities large, but nonschool factors continue to influence children as they progress through school.

So, while the "new accountability" focus on student achievement was a step forward in some ways, both state and federal policies misuse test scores in ways that violate the Fundamental Principle of Accountability. These misuses have consequences. Performance measures based on student test score attainment are partly responsible for pressuring schools to exclude students from testing (Figlio 2005) and pushing out good teachers (Cloftel et al. 2004) who are frustrated by a system that punishes them no matter how well they perform. It's also well known that school systems, especially urban school districts, "spin their wheels" in an endless cycle of reform and changes in curriculum, instruction, and leadership (Hess 1998). A continual cycle of reform might make sense if those schools were continually failing, but "low-performing" schools often aren't failing by more reasonable definitions. As I will show below, many low-attainment schools are actually high-performing, and the reverse is also true. While we have focused on low-attainment schools, high-attainment schools present an equally large, but hidden, problem. In these schools, students enter with high scores, so their attainment remains high almost no matter what the schools do. These schools often largely ignore external accountability.

Fortunately, we can take steps to help avoid confounding out-of-school influences with school or teacher performance.

The Basics of Value-Added Accountability

Performance can be reasonably defined, according to the Fundamental Principle of Accountability, as what each school and teacher contributes to student outcomes. But how can we measure that?

One simple way to measure performance is by subtracting the initial level of student achievement test score from the end-of-year test score. This is illustrated in Figure 1, in which each arrow represents the achievement growth of students in two hypothetical schools: Rockefeller Elementary and Smith...
Elementary. The bracket in the lower left corner indicates the differences in initial achievement — the "starting gate inequality" — and shows that Rockefeller students begin in a stronger position than Smith students. Growth measures allow us to compare schools after subtracting the starting gate inequality.

Figure 1 shows that, despite starting differences in attainment, the trajectory of growth is the same. Examining attainment and growth yield different conclusions about performance. Given that the starting gate inequalities are outside of school control, growth is a more fair way to compare than attainment. If we judged these schools on student attainment, Rockefeller would be the clear winner. Accountability systems have been doing just that for decades. We shouldn’t be surprised that low-income schools have complained vigorously about the lack of fairness in attainment measures.

By focusing on growth, we’re really coming up with a more reasonable comparison. Here’s another way to think about it: Growth measures allow us to make better predictions of future student achievement. What’s the best way to predict how students will do next year? The simple answer is to look at what students did last year. But predicting growth based on the prior test score alone is somewhat simplistic because some school factors that affect growth are still outside educators’ control. At the school level, for example, schools have little authority over their budgets, their curricula, or even, in many districts, teacher selection.

"Value-added" refers to statistical techniques that make predictions based on information that includes but goes beyond prior student achievement. Once predictions are made, they’re compared with what actually happens. Schools whose students reach higher achievement levels than predicted have high value-added scores. Conversely, schools whose students end up with scores lower than predicted have low value-added scores.

You might think that shifting from attainment to value-added models is irrelevant because both are based on the same student test scores. But there is only a modest relationship between attainment and value-added models (Weiss 2008). To put this in perspective, suppose we labeled each of 20 schools as either "effective" or "ineffective" based on attainment measures and then relabeled the schools using value-added measures. If there was no relationship (or "correlation") between attainment and value-added measures, then 10 schools would have the same label under both measures and the other 10 schools would have opposite labels (for example, "effective" according to the attainment measure but "ineffective" according to the value-added measure). In that case, the attainment would be no better than guessing with the flip of a coin.

According to Weiss, it turns out that 14 schools, rather than 10, would end up with the same label under both measures, and six schools would have different labels. This means, for example, that out of every 20 schools, three high-attainment-score schools would be mislabeled as high-performing and three low-attainment schools would be mislabeled as low-performing. This is why the attainment-based "adequate yearly progress"

If the question is whether teacher value-added is better than the existing system of evaluation, then the answer is probably "yes."

Much of the discussion here is based on the author’s forthcoming book about value-added for educators and policy makers (Harvard Education Press, in press).

REFERENCES


and “growth-to-proficiency” adopted by the federal government both violate the Fundamental Principle of Accountability.

Strengths and Weaknesses

No performance measure is perfect. An important part of measurement development is identifying the nature and severity of potential errors. Two general types of errors are always present to some degree — systematic error and random error.

With systematic error, we’re more likely to make a mistake with certain types of teachers and schools. I discussed one important example of systematic error above — the tendency of low-attainment-score schools to be mislabeled as low-performing. In this sense, snapshot-based school performance measures — such as percent of students passing a standard of proficiency — are systematically biased against schools serving low-attainment-score students.

But value-added measures don’t eliminate systematic error. For example, a spring test score may not be a good measure of what students know the next fall because many students lose what they’ve learned during the summer months.

This “summer learning loss” may affect some schools more than others. This means that students living in neighborhoods with fewer libraries and summer reading activities are likely to have more summer reading loss, and schools serving these students will be systematically disadvantaged.

Random errors are another story. They have to do with chance. Random errors are equally likely for any teacher or school. Suppose that we flip a coin and one person calls heads, but the coin comes up tails. This is not a systematic error because the error occurs for everyone with equal likelihood. The coin flip is random, and everyone has a 50-50 chance of calling it right. Random errors might seem more innocuous because they’re equally likely to arise with all teachers. But they can be just as problematic because these errors have the same effect of calling into question the conclusions we wish to draw about performance in the process of making high-stakes decisions.

Random error means value-added measures can be unstable over time — because the error is random, it changes over time for each individual teacher. Koedel and Betts (2007) found that only 35% of teachers ranked in the top fifth on teacher value-added measures one year were still ranked in the top-fifth in the subsequent year. This suggests that the other 65% of high-performing teachers actually got worse relative to their peers over a short period of time — some dramatically worse. McCaffrey and his colleagues (2009) show that results are more stable over time when using more years of data to evaluate each teacher. Therefore, at the very least, policy makers should avoid using teacher value-added measures based on only one or two years of information.

Many critics of value-added accountability also raise significant and valid concerns about the achievement tests underlying value-added measures. In addition to the test content and scaling procedures, there are questions about the timing of test administration and students who switch schools mid-year. All of these problems contribute to systematic and random errors. Another limitation — that tests aren’t available in all subjects and grades — means that value-added accountability can be applied only to a small percentage of teachers. This isn’t ideal, though teachers in different subjects and grades will always be evaluated differently, no matter what type of evaluation system is used.

Again, no performance measure is perfect, and value-added measures are no exception. Random and systematic errors reduce the accuracy of these measures and make it less likely that our conclusions are accurate. This, in turn, affects how we use value-added analyses in education policy.

Using Value-Added Measures

A basic principle of measurement is that the validity of a measure depends on what conclusion one is trying to draw from it. Value-added measures are good for some purposes and not for others.

School vs. Teacher Value-Added Analyses: School value-added analyses are less controversial than teacher value-added measurement. The reasons for this are clear. We already use school-level performance measures based on attainment and are likely to continue to do so. Value-added measures of school performance are clearly superior to attainment measures in order to maximize fairness.

The situation is more uncertain with teacher value-added measures. On the one hand, teacher value-added measures are unstable and probably don’t fully account for student tracking. On the other hand, teachers are arguably the most important school resource,
but teacher performance appears to vary widely (Sanders and Horn 1998; Rivkin, Hanushek, and Kain 2005) in ways unrelated to certification (Goldhaber and Brewer 2000) and preparation (Harris and Sass 2007). There is also wide agreement that the current teacher evaluation systems are broken. If the question is whether teacher value-added accountability is better than the existing system of evaluation, then the answer is probably “yes.” Less clear is whether teacher value-added accountability would be better than other feasible policy alternatives, such as improved principal and peer assessments.

A middle option would be evaluating teams of teachers by grade and subject area (math or reading). This is an attractive option for many reasons: It would reduce systematic and random errors because it includes more students and avoids tracking concerns. Team measures are less threatening to individual teachers. They also could do more to facilitate cooperation and coordination among teachers within each team.

Cooperation could also be undermined if teachers (or teams) are compared with others within a school or small district. This too can be avoided by comparing each teacher (or team) to those across an entire state. No single teacher would have any influence on how another teacher is rated.

Low-Stakes vs. High-Stakes. Many value-added discussions assume that the measures will be used for teacher merit pay, tenure, and other high-stakes decisions. These are possibilities, but not the only ones. The least controversial use of value-added measurement is for program evaluation. Indeed, many of the studies cited above do exactly that. They don’t report value-added scores for any individual teacher or school, but simply look for patterns across large numbers of teachers and schools to learn, for example, about the effects of professional development programs or how well teacher certification distinguished effective from ineffective teachers.

Other options include calculating value-added scores for individual teachers and providing this information privately to teachers and their principals, without attaching high stakes. Another small step would be using the information for creating professional development plans, including plans for individual teachers.

Formative vs. Summative. Value-added measures provide summative assessments of teacher performance—they indicate whether teachers are doing well in terms of one important student outcome. But value-added measurement is often criticized for not providing information about what steps can be taken to improve teaching. Value-added measures can be made more “formative” in this sense by estimating value-added by subject or even by specific domains in subjects. But the more important response is that no single measure is likely to fulfill both the formative and summative functions very well.

For this reason, any use of value-added accountability, especially for individual teachers, should be coupled with observations by school principals or neutral “peer” assessors. These additional performance measures not only could provide more formative information to help teachers take concrete steps forward, but they could reduce systematic and random errors. Such a combination represents the application of a long-standing rule of thumb in assessment-based decision making: When making decisions of consequence, more information is generally better.

Again, the issue is whether to use value-added measurement, but how.

The Case for Experimentation

Two things are clear from this discussion: First, school value-added models are better than the alternatives at the school level. Policy makers have been too willing to accept the conclusion that low-attainment schools are also low-performing, and this been devastating for high-poverty schools. These schools spin their wheels even more and have an even harder time attracting and retaining effective teachers.

Second, the current system of teacher evaluation and accountability has enough major problems that it’s worth experimenting with some uses of value-added assessment of teacher effectiveness. Any initiatives should be implemented as part of a holistic system that includes other measures. Evaluating the effects of these measures is important because we know little about how and whether value-added accountability improves how schools work.

Value-added measures have potential, but we can’t lose sight of the larger purpose: measuring performance in a way that facilitates genuine accountability, a sense of mission, and sound messages. If value-added accountability can do these things even a little, there’s little doubt that they’ll improve schools and help students.