Performance or Effectiveness? A Critical Distinction for Teacher Evaluation

By Justin Baeder on October 28, 2011 11:13 AM | No recommendations

*Guest post from **Rod McCloy** & **Andrea Sinclair***

Teacher evaluation has become a major focus of reform at the highest levels of education policymaking. The Obama administration awarded states more points for plans to improve teacher evaluation in their Race to the Top applications than for nearly any other policy area. The administration's Blueprint for Reform for reauthorizing the Elementary and Secondary Education Act (i.e., NCLB) would require states to revamp teacher evaluation to receive significant amounts of federal funding. The administration has also allocated federal School Improvement Grant (SIG) money for "persistently low-performing" schools adopting the Transformation model, which requires an overhaul of current teacher evaluation practices. And now, most recently, to receive a waiver from the cornerstone requirement of NCLB—that all students be proficient in math and language arts by 2014—states must create new teacher evaluation guidelines. In all of these instances, it is required that the teacher evaluation system be revamped to include student achievement as a significant component.

Industrial-organizational (I-O) psychologists have much to offer this discussion. I-O psychologists apply psychological principles in an attempt to understand, predict, and improve workplace behavior. They touch on many topics, including personnel selection, training evaluation, job design, and organizational development. Arguably, though, the key concern for I-O psychologists is how to measure, predict, and improve job performance. We therefore are acutely aware of many problems associated with measuring individuals' work performance. Even so, our field wandered about for decades using **myriad measures as performance indicators**, propagating substantial confusion in the research literature. In the early 1990s, however, a theory of job performance was proposed (Campbell, McCloy, Oppler, & Sager, 1993) that defines performance in a meaningful, explicit way and provides guidance to researchers and practitioners needing to choose performance measures for use in their studies and real-world settings, respectively.

The theory defines performance as

> ". . . synonymous with behavior. It is something that people actually do and can be observed. By definition, it includes only those actions or behaviors that are relevant to the organization's goals and that can be scaled (measured) in terms of each individual's proficiency (that is, level of contribution). Performance is what the organization hires one to do, and do well. Performance is not the consequence or result of action, it is the action itself... [and] consists of goal-relevant actions that are under the control of the individual" (Campbell et al., 1993, p. 40, emphasis added).

Performance on any job is complex; that is, it is not just "one thing" but instead consists of multiple, distinguishable components (it is multidimensional). One result of the complexity of job performance is that the notion of "overall job performance" is often not a meaningful concept. It is preferable to measure the components of performance separately. Decision-makers often invoke "overall job performance" because they need a single score upon which to base their decisions, but practical demands do not eradicate scientific reality.

The theory also makes a critical distinction between performance and three other concepts: effectiveness, productivity, and utility. Effectiveness is defined as

> ". . . the evaluation of the results of performance. By definition . . . a measure of effectiveness is controlled by more than the actions of the individual. Dollar amount of sales is an obvious example" (Campbell et al., p. 41).

Here, we see the major problem with teacher performance being defined as a function of students' standardized test scores. Rather than indexing teacher performance (behaviors under the teacher's control), such a measure is an indicator of teacher effectiveness (the results of performing/not performing those behaviors).

To highlight the difference between performance and effectiveness, consider what it takes to be a high-performing sunglasses salesperson. The required characteristics are those that any good salesperson should possess: knowledge of the product line, an outgoing and friendly demeanor, and excellent interpersonal (e.g., instructing, social perceptiveness) and communication skills and abilities (e.g., active listening, oral expression). Possessing all of these characteristics, however, will NOT guarantee effectiveness, nor will their absence guarantee ineffectiveness! This is because effectiveness depends on factors extraneous to the person's behavior, such as location (selling sunglasses in Binghamton/Seattle as opposed to Miami/Las Vegas). Imagine that we know the Seattle salesperson to be a much higher performer (better knowledge of the product line, more outgoing and friendly, etc.) than the Miami salesperson. Despite this, the Seattle salesperson yields lower sales than the Miami salesperson. Thus, the Seattle salesperson is a better performing—but less effective—salesperson.

What does this mean in terms of personnel decisions? Should we replace the Seattle salesperson in hopes of improving sales—perhaps by sending the Miami salesperson to Seattle? On the contrary, if anything, we should replace the low-performing Miami salesperson. Just imagine how effective a good performer would be in the desirable Miami sales market!

One should not infer from this example that we consider effectiveness to be unimportant. Indeed, it is the bottom line for most organizations. Nevertheless, it is at the level of performance that organizations have the greater capacity to influence results with appropriate interventions. The critical point is that we should not equate performance with effectiveness. This is not simply an exercise in semantics. They are separate, distinct concepts. Each is important, but they tend to be discussed interchangeably—to everyone's detriment.

So what are the ramifications of the performance/effectiveness distinction for teacher evaluation? This will be the topic of the next blog post.

Dr. Rodney A. McCloy *is a Principal Staff Scientist for the* ***Human Resources Research Organization (HumRRO)****. With more than 20 years of experience conducting and directing personnel research, he serves as an in-house technical expert and a mentor to junior staff. His assessment and testing experience has spanned both cognitive and non-cognitive domains and has involved several large-scale assessment programs (Armed Services Vocational Aptitude Battery, National Assessment of Educational Progress, General Aptitude Test Battery). He has served as adjunct faculty at both The George Washington University and George Mason University. He is a Fellow of the American Psychological Association (APA) and the Society for Industrial and Organizational Psychology (SIOP). He received his Ph.D. in Industrial-Organizational Psychology from the University of Minnesota in 1990.*

Dr. Andrea L. Sinclair *is a Senior Scientist in* ***HumRRO****'s Validity Investigations for Education and the Workplace (VIEW) Program. She conducts research in education, government, military, and private*

sector settings with a particular focus on performance measurement and program evaluation. She regularly develops performance measurement instruments, surveys, and observation and interview protocols for use in schools. In addition, she regularly advises clients on the validity and reliability of their assessment systems and on the development of competency models. She received her Ph.D. in Industrial-Organizational Psychology from Virginia Tech in 2003.

Reference:

Campbell, J.P., McCloy, R.A., Oppler, S.H., & Sager, C.E. (1993). A theory of performance. In N. Schmitt & W. Borman (Eds.), Personnel selection in organizations (pp. 35-70). San Francisco, CA: Jossey-Bass.

"Job Performance" Measures from the Research Literature (from Campbell, McCloy, Oppler, & Sager, 1993, p. 36):

- Time to complete a training course
- Grades or achievement test scores earned in training
- Number of errors made in a simulator
- Number of Tinkertoy figures assembled in a 45-minute experimental session
- Number of one-minute marketing interviews completed outside a shopping center in one day
- Number of pieces produced
- Number of defective pieces produced
- The total or average cost of the pieces produced
- Number of proposals written
- Total value of contracts won
- Total value of sales
- Number of grievances or complaints incurred
- Length of tenure in the organization
- Total days absent
- Salary level
- Promotion rate within an organization
- Percentage over budget
- Supervisor, peer, subordinate, or self ratings of "overall" performance
- Scores on a paper-and-pencil job knowledge test
- Scores on a professional certification test
- Number of citations in the citation index over a 3-year period
- Promotion rate within an organization
- Number of refereed journal articles published in a 6-year period

Categories: **Performance**

Ramifications of the Performance/Effectiveness Distinction for Teacher Evaluation

By Justin Baeder on November 4, 2011 2:07 AM | Recommend

*Guest post from **Rod McCloy** & **Andrea Sinclair***

In our **initial blog entry**, we argued that it is essential to differentiate performance (behaviors people engage in on the job; i.e., what people do) from effectiveness (the results of performance) when conducting teacher evaluation.

In this entry, we discuss how doing so can clarify the discussion surrounding teacher evaluation.

We first must specify just what it is we intend to evaluate: performance? effectiveness? something else? It is critical that we answer this question clearly, because performance and effectiveness are different criteria determined by different variables, which suggests the potential for different interventions to improve them.

Recalling our sunglasses salespersons from **our initial blog entry**, we believe that the performance/effectiveness distinction has several implications for the classroom:

- The best-performing teachers will not necessarily be the most effective teachers (and vice versa);
- Placing effective teachers from one setting into a markedly different setting (e.g., moving highly effective teachers from a suburban school to a low-performing urban school) could lead to disappointing outcomes;
- By focusing on teacher performance, we can maximize teacher effectiveness as a by-product (recall the suggestion to move the higher performing but less effective salesperson from Seattle to Miami).

Indeed, measuring teacher performance (practices/behaviors teachers engage in) gives us the best chance of providing (a) teachers with useful developmental feedback on their practices and (b) educators/administrators with input on teacher training programs.

Most current initiatives require schools to include students' standardized test scores or academic achievement (indices of effectiveness rather than performance) in their teacher evaluations. For example, in states with winning applications for Race to the Top grants, this student information must constitute at least 50% of the overall teacher evaluation.

Those using such information in their teacher evaluations need to be cautious about the attributions they make based on such data. In accordance with the performance theory (Campbell, McCloy, Oppler, & Sager, 1993), one should not attempt to identify teacher-level interventions or make judgments about teacher performance by examining outcomes contaminated by influences beyond the teacher's control. Although student achievement data could alert evaluators as to when they should look more closely at a teacher's performance ratings to help determine if there is something about the teacher's performance that contributed to student achievement, teacher performance measures are required to help identify the types of behaviors that might need to improve. Again, effectiveness is important and useful in its own right, but it is not the same thing as performance and should be kept distinct.

You might be asking, "But what about Value-Added Modeling (VAM)? It purports to isolate teacher impact on student performance by statistically controlling for external influences. Doesn't this mean that VAM provides information about teacher performance?" To our minds, there are at least two shortcomings of VAM with regard to teacher performance. First, VAM is at best an indirect means of obtaining information about teacher performance. We believe it preferable to define performance explicitly, rather than taking performance to be the residual of a subtractive process via statistical control of certain select "other factors."

If you want to measure teacher performance, then measure it directly. Doing so will force you to delineate the behaviors of interest (i.e., what you define performance to be) and increase your chances of identifying promising interventions for improving performance (and, thereby, effectiveness). Second, VAM seems to limit the definitions of both teacher effectiveness (to students' test scores) and teacher performance (to only those behaviors that increase student achievement on tests, and this assumes that we know which behaviors those are). Thus, both effectiveness and performance as defined by VAM are likely deficient concepts.

Please do not let our concerns regarding VAM lead you to believe we are anti-testing. On the contrary, we are staunch supporters of standardized testing. Nevertheless, current VAM seems to discount the inherent complexity of teacher performance and teacher effectiveness, artificially constraining their definitions and indicators. We enthusiastically endorse the use of empirical data, but convenience (students' test scores are available and standardized, at least within states) must not trump relevance (students' test scores tell us little about specific teacher behaviors) when choosing data to serve as the foundation for high-stakes personnel decisions.

Our next entry will present our recommendations for developing teacher evaluation systems.

Dr. Rodney A. McCloy *is a Principal Staff Scientist for the* ***Human Resources Research Organization (HumRRO)****. With more than 20 years of experience conducting and directing personnel research, he serves as an in-house technical expert and a mentor to junior staff. His assessment and testing experience has spanned both cognitive and non-cognitive domains and has involved several large -scale assessment programs (Armed Services Vocational Aptitude Battery, National Assessment of Educational Progress, General Aptitude Test Battery). He has served as adjunct faculty at both The George Washington University and George Mason University. He is a Fellow of the American Psychological Association (APA) and the Society for Industrial and Organizational Psychology (SIOP). He received his Ph.D. in Industrial-Organizational Psychology from the University of Minnesota in 1990.*

Dr. Andrea L. Sinclair *is a Senior Scientist in* ***HumRRO****'s Validity Investigations for Education and the Workplace (VIEW) Program. She conducts research in education, government, military, and private sector settings with a particular focus on performance measurement and program evaluation. She regularly develops performance measurement instruments, surveys, and observation and interview protocols for use in schools. In addition, she regularly advises clients on the validity and reliability of their assessment systems and on the development of competency models. She received her Ph.D. in Industrial-Organizational Psychology from Virginia Tech in 2003.*

Reference:

Campbell, J.P., McCloy, R.A., Oppler, S.H., & Sager, C.E. (1993). A theory of performance. In N. Schmitt & W. Borman (Eds.), Personnel selection in organizations (pp. 35-70). San Francisco, CA: Jossey-Bass.

Categories: **Evaluation**,**Performance**

Recommendations for Developing Teacher Evaluation Systems

By Justin Baeder on November 20, 2011 8:01 PM | **2 Recommendations**

*Guest post from **Rod McCloy** & **Andrea Sinclair***

In this blog entry, we provide some of our recommendations for developing teacher evaluation systems. These recommendations rest on the performance theory (Campbell, McCloy, Oppler, & Sager, 1993) presented in our previous two posts ("**Performance or Effectiveness? A Critical Distinction for Teacher Evaluation**", "**Ramifications of the Performance/Effectiveness Distinction for Teacher Evaluation**") and its differentiation between performance and effectiveness.

Our recommendations for developing teacher evaluation systems then are as follows:

1. Develop appropriate performance measures, keeping multidimensionality (multiple distinguishable components of teaching) in mind. Performance measures should focus on those behaviors teachers are hired to do and do well. Also, teacher performance is complex. We need to look for these various dimensions of performance rather than settling for an "overall performance" measure.

2. Maintain the performance/effectiveness distinction. This distinction is essential and not just an exercise in semantics. Performance drives effectiveness, but effectiveness regards the results of that performance; they are not the same thing. Keeping these concepts distinct allows us to learn about both; confounding them prohibits us from learning about either one.

3. Focus teacher evaluation on performance rather than effectiveness. This means that evaluators will focus on behaviors under each teacher's control, thus facilitating the identification of appropriate developmental interventions for improving teacher behaviors and greatly improving the perceived fairness and usefulness of the evaluation system.

4. Consider teaching to be not just one job but possibly several jobs. Teachers in different settings (e.g., a low-performing school with poorly motivated, low-achieving students vs. a high-performing school with highly motivated, high-achieving students) will likely need to enact different strategies and engage in different behaviors to be effective. Therefore, the "teacher" occupation might be profitably viewed as comprising several jobs, each with its own set of performance dimensions that likely differ across contexts/settings. It seems to be explicit in teacher training that the job of a primary school teacher differs in meaningful ways from that of a secondary or post-secondary school teacher. It might be just as useful to consider the possibility that the various environmental settings in which teachers find themselves mandate different sets of work behaviors that should be measured and perhaps even different means by which teachers can be maximally effective.

Measuring teacher performance is challenging. Many important behaviors are likely difficult to observe and even more difficult to accurately measure. Nevertheless, industrial-organizational (I-O) psychologists know how to define and measure performance in jobs where behaviors can be difficult to observe (e.g., managers). Perhaps we can work together to improve the evaluation process for one of society's most critical professions.

***Dr. Rodney A. McCloy** is a Principal Staff Scientist for the **Human Resources Research Organization (HumRRO)**. With more than 20 years of experience conducting and directing personnel research, he serves as an in-house technical expert and a mentor to junior staff. His assessment and*

testing experience has spanned both cognitive and non-cognitive domains and has involved several large-scale assessment programs (Armed Services Vocational Aptitude Battery, National Assessment of Educational Progress, General Aptitude Test Battery). He has served as adjunct faculty at both The George Washington University and George Mason University. He is a Fellow of the American Psychological Association (APA) and the Society for Industrial and Organizational Psychology (SIOP). He received his Ph.D. in Industrial-Organizational Psychology from the University of Minnesota in 1990.

Dr. Andrea L. Sinclair *is a Senior Scientist in* ***HumRRO****'s Validity Investigations for Education and the Workplace (VIEW) Program. She conducts research in education, government, military, and private sector settings with a particular focus on performance measurement and program evaluation. She regularly develops performance measurement instruments, surveys, and observation and interview protocols for use in schools. In addition, she regularly advises clients on the validity and reliability of their assessment systems and on the development of competency models. She received her Ph.D. in Industrial-Organizational Psychology from Virginia Tech in 2003.*

Reference:

Campbell, J.P., McCloy, R.A., Oppler, S.H., & Sager, C.E. (1993). A theory of performance. In N. Schmitt & W. Borman (Eds.), Personnel selection in organizations (pp. 35-70). San Francisco, CA: Jossey-Bass.

Categories: **Evaluation**,**Labor Market & Workforce**,**Performance**