Introduction

Decades ago, I directed a Center for Performance Assessment at the Northwest Regional Educational Laboratory in Portland, OR. Performance assessment is a method of educational measurement in which the evaluator observes the performer's behavior or products in order to make a subjective professional judgment about the level of proficiency demonstrated. Each year, my team and I would select one prominent application of this method to be the focus of our research and development, such as writing assessment, speaking assessment, science lab, and so on. Our mission was to improve our understanding of the methodology so we could train users more effectively.

One year, we selected teacher evaluation as our focus—the observation and judgment of teachers' pedagogical skills by their supervisors. While our studies always tried to address both formative and summative uses of performance assessment, in this case, we knew that we were dealing with a summative application. At that time, references to the formative side of teacher evaluation were rare in the professional literature, as well as in local policy and practice.

We found a lack of quality in this particular application of the performance assessment methodology. Often, the performance criteria upon which teachers were to be judged were vague. Typically, raters were untrained in the data collection and evaluation process. The sample of performance gathered typically was far too thin (e.g., one observation per year) to support any inferences about the efficacy of a teacher's

instructional activities. Feedback was superficial if there was any at all, and the consequences were virtually always the same: retention. All of this has been well documented in the professional literature.

As a sidebar, we launched a specific and wide-ranging search for teachers who had experienced important professional growth as a direct result of the evaluation of their performance by their supervisor. We found almost none and the few we did identify all had been experienced under the same supervisor. We studied those events and that principal and derived a set of keys to successful formative teacher evaluation in our final report (Stiggins and Duke, 1988).

Upon completion of that work, I left teacher evaluation behind and returned to my career-long passion: understanding and improving teachers' day-to-day classroom assessments of student achievement.

Now, all of these years later, I am compelled to return to the practice of teacher evaluation. I am lured by the very exciting, but the unprecedented and very strong, national movement to weave student achievement into the teacher evaluation process. It has always been my belief that we should consider this dimension of teacher performance. At the same time, I have been stunned by the practice—required by policy—of defining student growth as changes in scores on annual standardized accountability test scores. We definitely should NOT be doing this. I am a classically trained specialist in educational measurement and was director of test development at ACT for five years before changing my focus to classroom assessment. I know of no one in the measurement community who thinks this is an appropriate use of annual standardized test scores. As the 2013 teacher strike unfolded in the Chicago Public Schools over the issue of the use of annual standardized test scores to evaluate teachers, I found myself appalled that these school leaders would propose (indeed, demand) this inappropriate application of these tests. I had hoped the teachers would not give in to such a misguided practice, but they capitulated and I immediately started work on this book.

I believe we can and should consider student growth in evaluating the efficacy of teachers' instructional interventions. But the preponderance of evidence of impact should come from another source: classroom assessments of student learning. This book is about why. However, certain conditions must be satisfied in schools and classrooms in order for this source of evidence of teacher impact on learning to be tapped. Those conditions are easily satisfied. My presentation also is about how this can be done.

Chapter 1 explores the change in school mission that has opened the door to the consideration of student growth in the evaluation process. It sheds light on the origins of and problems created by federal and state policies requiring standardized test scores for teacher evaluation. In addition, I outline the keys to the truly effective evaluation of teacher performance, including the criteria that must be satisfied for the appropriate consideration of growth in student achievement.

Chapter 2 analyzes the specific problems associated with the consideration of annual accountability test scores, centering predominantly on the danger of a mismatch between the learning targets tested and the targets that make up any individual teacher's specific instructional responsibilities. I also confront the popular practice of detecting student growth using value-added analysis models. In this context, such analysis cannot support inferences about the impact of a teacher's classroom interventions.

In Chapter 3, I describe an alternative approach to the consideration of student growth in the teacher evaluation process that is both practical and technically defensible under the right circumstances. Indeed, this approach promises not only fair evaluation for teachers but very strong learning benefits for them and their students. It relies on evidence of student growth derived from classroom assessments; however, for this option to be implemented effectively, all involved must bring to the process a sufficiently high level of assessment literacy (an understanding of the basic principles of sound assessment practice).

Chapter 4 addresses a currently popular approximation of a classroom assessment-based definition of student growth I am advocating called Student Learning Objectives. I will analyze the similarities and then detail the problems with this version that are remedied with the model defined in Chapter 3.

Chapter 5 describes the levels of assessment literacy needed in every classroom and school in order for teachers and supervisors to take full advantage of my approach. The challenge we face is that pathways for attaining those levels of competence have almost never been open to teachers or school leaders. The time has come to put a solid foundation of assessment literacy in place. As I will explain, sufficiently high levels of competence in classroom assessment can be achieved very economically.

Faculty-wide assessment literacy is not sufficient in and of itself, however. It represents only one part of a multi-layered assessment system infrastructure that local district and school leaders must put in place to achieve truly effective teacher evaluation. In Chapter 6, I identify a series of essential assessment actions local leaders must take to build the local assessment systems that can support effective teacher evaluation and meet the wide-ranging informational needs of practitioners in their organizations. The benefits associated with the completion of these actions are legion and, as I will demonstrate in this concluding chapter, can include immense learning benefits for students.

A Message for All of the Key Audiences

For *local, state, and federal policy makers*, the message implicit in my presentation is that much of what has been unfolding in recent years by way of political action surrounding the consideration of student growth in teacher evaluation is technically unsupportable and, therefore, wrong. I will defend this proposition. Currently required practices very often are doomed to ineffectiveness for a variety of practical reasons. I will detail those. As policy makers, you must

develop sufficient understanding of the differences between sound and unsound assessment and teacher evaluation practices in order to conduct a thoughtful reconsideration of your decisions. There are far more effective ways to consider student achievement in the evaluation of teachers than emerging laws and policies reflect. They deserve a fair trial.

School leaders at state, local district, and building levels are in charge of the teacher evaluation process and so are key players in this story. Almost all of you come to this challenge without having been given the opportunity to understand the basic principles of sound assessment practice. That professional development is readily available to you. My hope is that my presentation will trigger assertive action on your part to delve into it and see to the development of an essential foundation of assessment literacy throughout your organizations. Then and only then can you (a) provide the policy makers around you with the guidance they need to drive sound practice, and (b) lead the kind of faculty development needed to build a proper teacher evaluation system.

Historically, *teachers* have been denied the benefits of high-quality performance evaluation practices. Those practices have offered little by way of assistance in promoting your professional development and have been so locked up procedurally as to lead to few positive impacts in terms of personnel action. The strategy I offer for consideration of student achievement in the evaluation of your performance can provide positive and productive assistance on both fronts. As you will see, it affords each teacher the opportunity to play a role in generating focused evidence on student learning that can improve your instruction. In addition, it places you in the heart of the summative evaluation process by assigning you primary responsibility for building and presenting your own case for the efficacy of your teaching practices and skills. In any job evaluation process, one can ask for little more.