Florida Association of Teacher Educators Journal Volume 1 Number 4 2004 59-66
http://www.fate1.org/journals/2004/lang.pdf

Measuring Teachers' Effects on K-12 Learning

W. Steve Lang
University of Florida at St. Petersburg, St. Petersburg, Florida

Judy R. Wilkerson
University of Florida at St. Petersburg, St. Petersburg, Florida

Accreditation and program approval requirements nationwide require the assessment of teachers' knowledge, skills, and dispositions, as defined in national and state standards of the profession. They continue to require that the result of teachers' instruction on children's learning be assessed, as well, so that all children learn. It is this bottom line result that is the focus of this paper, which explores implementation strategies for measuring the impact or effect on student learning that teachers have. Stated in simpler terms, this article discusses measures of teacher effectiveness that track systematically K-12 student progress as part of a larger assessment system that looks at all four elements – knowledge, skills, dispositions, and impact.

Measuring Teachers' Effects on K-12 Learning

How Can the Effect on K-12 Learning Be Demonstrated?

Just as is the case in measuring knowledge, skills, and dispositions, a variety of assessment strategies used in multiple assessment types is necessary. Making a decision about whether or not a teacher can have a positive effect on student learning cannot be made from one piece of evidence any more than we can make a decision about pedagogical skill based on one measure. As important as that internship observation form is, there is more to assessing good teaching than observing performance skills for a 30-minute period! The best assessments link the teacher performance on KSD's directly to students' progress, performance, or achievement.

A variety of standards-based assessments that are hierarchical in nature and tap a variety of attributes are advisable. The use of increasing levels of inference, or judgment, increases the confidence we can have in the decision about whether or not teachers impact K-12 learning (or have pedagogical skills or attitudes necessary to do so). Let us begin this discussion with multiple measures and inference and why they are so important.

Multiple Measures and Levels of Inference

A combination of assessment strategies allows for increasing levels of inference on the part of both candidates and faculty, thereby increasing the potential validity of the decisions made about what candidates can do. While we strongly acknowledge the successful use of gain score methodologies such as the ones currently accomplished so well in the TWSM and Renaissance models, it is but one measure of impact and has some serious weaknesses. Achieving a gain in scores on the same test after instruction is less difficult than teaching the students to produce a complex product requiring that they synthesize material learned into some new result. It is also easier to score the tests and compute the mathematics to determine who gained how much, since objective tests usually have correct/incorrect answers that can be counted. Products requiring synthesis of factual knowledge into some new work are more subjective, requiring much more professional judgment. Two examples of increasing levels of inference follow.

- It is easier for students to answer (and teachers to score) multiple choice questions requiring factual knowledge about weather patterns than it is for students to create (and teachers to score) the family hurricane evacuation plan.

- Sorting materials into biodegradable versus non-biodegradable categories on a matching test is easier than determining the level of commitment to recycling in the family or community

Projects such as these latter two examples (the evacuation plan and the recycling survey) are not subject to gain scores but do provide evidence of more in-depth learning and exhibit a higher level of critical and creative thinking skills than do the pre-post-test design referenced first in each bulleted example. The critical point here is that multiple measures, with increasing levels of *difficulty and inference*, add richness to the learning and assessment process and improve the validity of decisions about teacher ability to influence student learning.

The Assessment Problems with Current Systems

Nationally, there appears to be an over-reliance on the pre-test/post-test model of gain scores that makes the measurement of some more sophisticated learning problematic. Besides the threats to validity (Campbell & Stanley, 1963) that we all learn in beginning research courses, this model employs an objective (traditional) paper and pencil test. While objective tests can be written to include higher order thinking skills or delivered on computer, it is both difficult and atypical to assess more than a subset of knowledge at a lower cognitive level. Most teachers have trouble writing anything but memorization or simple application/analysis level questions. For that reason, these models that depend on standardized objective tests usually restrict our ability to look at progress in more complex skills such as evaluation and synthesis. It is also something that cannot be done realistically in practice for every unit taught without giving much more classroom time to assessment. Finally, there are serious job-relatedness issues with

this approach. Job-relatedness is one of the most important elements to provide evidence of construct validity (Wilkerson and Lang, 2004).

If one turns to alternative assessments (e.g., projects, essays, or performances), for assessment of complex learning by students then the pre/post model with gain scores is no longer practical. Think about a unit on some aspect of politics. The students could have a pre-test on the facts, but wouldn't it be good to then have them do a debate or something that requires a higher level of thinking skill than just to repeat the test again so the teacher can calculate gain scores? So, we advise that the notion of gain scores be downplayed a little – not eliminated because it is certainly important in some circumstances – but downplayed. There is more to life than repeated measures. We think most people resort to gain scores on standardized tests out of lack of knowing how to do anything else. We prefer to imagine a "ruler" of ever increasing complexity and difficulty of ability or achievement.

Showcasing Student Work Rather Than Teacher Work

An example of how to build confidence would be to create a mini-portfolio of teacher candidate work, all of which focuses on increasingly difficult and varied tasks that show some level of impact on what students can do as a result of the teacher's work with them. In Figure 1, we provide an example of such a set of tasks that can be combined into a portfolio of this type. The tasks are correlated with the Florida Educator Accomplished Practices and briefly described in the figure. All of the tasks are drawn from a system developed for the Florida Alternative Certification Program Assessment System (Wilkerson, et al., 2002), now used by about 45 of the 68 school districts in Florida.

Figure 1: Thematic Portfolio Tasks Names and Descriptions and Aligned FEAPs

Task #	Task Name	FEAP	FEAP Indicators	Brief Description
09A	Classroom Management System	9	9.03, 9.06	The teacher creates a system to manage the classroom. The system includes rules; expected behaviors; procedures; and organization of space, time, and materials. Students are involved in the planning process.
05A	A Demographic Study of Your Students and a Plan to Meet their Needs	5	5.02, 5.04	The teacher compiles descriptive group data about students in the school and about specific students in his/her classes, presenting school and class data in chart form. The teacher analyzes these data and, after consultation with an ESE specialist and counselor in the school, the teacher documents the adaptations that he/she will need to meet the individual needs of students.
07C	Student Attitudes about School and Learning	7	7.01, 7.03, 7.04, 7.05	The teacher administers a survey to students to determine their attitudes about self-concept, reading/learning, autonomy, environmental mastery, family, and school. The teacher analyzes the results of the survey and develops strategies to motivate and support students. The product includes the analysis and the strategies along with a report of their effectiveness.
05C	Individual Planning for Intervention	5	5.07, 5.08, 5.10, 5.12	The teacher participates in an Individual Education Plan (IEP), Academic Improvement Plan (AIP), or Family Services Plan (FSP) as part of a team conference aimed at the instructional planning or intervention for an ESE student, LEP student, or a student identified as needing intervention. The teacher participates in alternative strategy planning or recommendations for the student and takes responsibility for implementation of a portion of the IEP or student intervention plan, keeping a running record of the progress/success of the intervention.
05B	Documentation of Diversity Accommodations	5	5.04, 5.05, 5.06, 5.07, 5.09	The teacher documents the adaptations that he/she has made to meet the individual needs of students.

01D and 11D	Case Study of a Student Needing Assistance	1	1.01, 1.04, 1.08, 1.09, 1.10, 1.12, 1.13	The teacher works with a student, his/her parent/guardian, and colleagues throughout the semester to improve the student's performance. This could be the student identified in Task 05C. The teacher keeps records of family and colleague contact targeted at improving the student's performance. The teacher also assesses that student's performance, reports results to all involved parties, and maintains records of the results and the discussions/reports of performance.
		11	11.01, 1.02, 11.03, 1.08, 11.10	
01E	Demonstration of Positive Student Outcomes	1	1.01, 1.04, 1.07, 1.11	The teacher demonstrates a positive impact on student learning in a major specified unit using multiple measures to determine mastery of objectives by individual students and the class as a whole.
04F	Portfolio of Student Work	4	4.01, 4.02, 4.03, 4.08	The teacher creates a showcase portfolio of student work over a semester or year. This portfolio includes samples of work from students in the teacher's class who have exceeded and/or met expectations with regard to targets for critical, creative, or higher-level thinking and content mastery and samples from students who were a challenge to teach. The teacher explains why each sample was selected and provides copies of the prompts use to generate the work (tests, lesson plans, or task instructions.)
		8	8.01	

Conclusion

The assessment ideas presented in this paper provide a viable approach to measuring the impact of teachers on K-12 student learning in a variety of ways, with increasing levels of inference. They have documented evidence from school districts of validity because they are job-related and authentic (Wilkerson and Lang, 2003 and 2004). There may be many other ways to assess teacher impact, but the important point here is that multiple measures, with increasing levels of inference, all focused on children, need to be at the core of the assessment system.

References

Campbell, D. T. & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Rand McNally, Chicago.

Elliott, Emerson (in press). *Student Learning in NCATE Accreditation.* National Council for Accreditation of Teacher Education. Washington, D.C.

Lang, W.S. (2002). *Alternative Certification Program Alignment Tasks to Ed-Ventures and the Survival Guide Content Validity Review*. Report: Bureau of Teacher Certification, Florida Department of Education, Tallahassee, FL.

National Council for Accreditation of Teacher Education (1999). *Professional Standards for the Accreditation of Schools, Colleges, and Departments of Education.* Washington, D.C.: Author.

Wilkerson, J. R. & Lang, W. S. (2004). A standards-driven, task-based assessment approach for teacher credentialing with potential for college accreditation. *Practical Assessment, Research, and Evaluation, 9(12).*

Wilkerson, J. R. & Lang, W. S., (2003). *Florida's alternative certification program assessment system: analysis of district coordinators' validity questionnaire on assessment tasks.* Report: Bureau of Teacher Certification, Florida Department of Education, Tallahassee, FL.

Wilkerson, J.R., Lang, W.S., Hewitt, M., Egley, R., & Stoddard, K. (2002). *Florida's Alternative Certification Assessment System.* Bureau of Teacher Certification, Florida Department of Education, Tallahassee, FL.

Author Note

W. Steve Lang, Ph.D., is an Associate Professor of Educational Measurement and Research at University of Florida at St. Petersburg in St. Petersburg, Florida.

Judy R. Wilkerson, Ph.D., is an Associate Professor of Research and Assessment at University of Florida at St. Petersburg in St. Petersburg, Florida.

Article Citation

Lang, W. S., & Wilkerson, J. R. (2004). Measuring teachers' effects on K-12 learning. *Florida Association of Teacher Educators Journal*, 1(4), 59-66. Retrieved from http://www.fate1.org/journals/2004/lang.pdf