

The Teacher Work Sample as a Context for Research¹

Mark Girod and Mark Schalock
Western Oregon University

Nora Cohen
Oregon State University

Contact information:

Mark Girod, Ph.D.
Western Oregon University
345 N. Monmouth Ave.
Monmouth, OR 97361
Phone: (503) 838-8518
Email: girodm@wou.edu

¹Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, San Diego, CA January '06

Introduction

Teacher education is straining under the burden of scrutiny like never before. Accountability pressures highlight the role of teacher preparation as research demonstrates the role of teachers in student achievement (Darling-Hammond, 1997; Fetler, 1999; Sanders & Horn, 1994; Wenglinsky, 2002). Despite this, teacher education suffers a dearth of well-articulated theory and empirical evidence connecting teacher preparation, teaching practices, and P-12 student learning. Cochran-Smith (2005) describes the challenge:

To get from teacher education to impact on pupil's learning requires a chain of evidence with several critical links: empirical evidence demonstrating the link between teacher preparation programs and teacher candidates' learning, empirical evidence demonstrating the link between teacher candidates' learning and their practices in actual classrooms, and empirical evidence demonstrating the link between graduates' practices and what and how much their pupils' learn (p. 303).

This linkage between preparation, practices, and P-12 student learning is inordinately complex (Berliner, 2002) though several methodologies are being used to bring definition to the task including teacher work sample methodology (TWSM). TWSM is based on an empirically validated framework for helping teacher candidates connect their actions to the learning of P-12 students (Girod, 2002; Elliott, 2004; Schalock, 1998). TWSM originated at Western Oregon University during the 1970s and has been adapted by numerous state education agencies and teacher education programs including the Renaissance Partnership and the states of Oklahoma, Tennessee, Kentucky, and Louisiana. The current political climate with its attention to teacher and student accountability and the shift in schooling from a norm-referenced, textbook driven system to a learner-center, standards-based system has highlighted the need for a framework like work sampling that offers fodder for potential theoretical and empirical connections between preparation, teaching practices, and P-12 student learning. This throughway will serve as the organizing framework for the remainder of this paper and we examine major efforts occurring first within the state of Oregon and now, more recently, expanding beyond Oregon to the broader national scene.

1980s through 2003

Within this time span, the focus of work in Oregon at Western Oregon University and the Teaching Research Institute included: K-12 student learning data generated through TWSM; the valid and reliable use of scores coming from TWSM within a high stakes initial licensure decision; the tradeoffs/ balance that must be addressed to take full advantage of the strengths of the methodology as both an instructional tool and assessment tool; and, finally as a measurement tool within a longitudinal research project on early career elementary school teachers. The major research and development tasks during this time were two-phases of the Teacher Effectiveness Project (TEP).

Pre TEP-I (1987-1994)

While much of the early work around the methodology centered on its design and implementation across teacher preparation programs in Oregon in response to the 1987 changes in licensure laws in Oregon, from a research standpoint, work focused on the student learning and contextual data being generated through its use. A long standing and seemingly intractable issue at the time was the lack of a defensible outcome measure in conducting research on teacher preparation. From 1989 through 1994 work lead by the WOU School of Education Research Design Team and staffed by Teaching Research faculty with financial support through the Provost's Office. This group designed and tested measures, formulated questions of interest to the faculty, and designed studies to answer these questions.

During this time data from nearly 1000 student teachers and 10,000 K-12 students were collected, analyzed and reported back to School of Education faculty. Data on student learning outcomes, observational data from school and college supervisors, and detailed information about classroom contexts in which student teachers were teaching were summarized and presented to faculty several times a year to support reflection on program design and adequacy.

A significant amount of work went into developing a method to compare student learning results across student teachers. This focused on both employing a metric developed by the late Jason Millman and assessing the cognitive demands of the assessments used by student teachers. A major purpose of this effort was to move student teachers beyond simple low-level recognition and recall types of assessments to higher level assessments addressing the kinds of student outcomes being demanded in K-12 schools at the time.

By 1994, the methodology and the student learning data were at a point where it seemed appropriate to have an external review conducted. In October 1994 a national invitational conference was held to expose the methodology and resulting learning gains, as a means for connecting teaching and learning for licensure and research purposes to a thorough, objective external evaluation. Thirty plus leading scholars in the areas of evaluation, research, teacher preparation, measurement, and policy spent three days reviewing the methodology and findings to date (see Conference Proceedings for participants and outcomes - citation). The overall reaction from conference participants was that while the work to date was commendable and the findings promising, further work was needed, and external funding to support such work was critical.

TEP-I (1995–1998)

After the October 1994 Conference, private funding was secured to carry the work forward as recommended by conference participants. A National Advisory Panel was formed, lead originally, until his death by Jason Millman of Cornell University. Dan Stufflebeam chaired the panel after Dr. Millman's death. Other panelists included: Ray Pecheone, Peter Airasian, Herb Walberg, Russel French, David Mandel, Robert Mendro, and Ruth Mitchell.

The work of TEP-I centered around researching and developing the conditions necessary for the valid and reliable use of scores – both student learning and teacher performance - coming from

TWSM within a high stakes initial licensure decision making context. Five conditions were deemed as both necessary and sufficient for this to occur. First, the performance task students teachers were being asked to carry out within the work sample must be valid. Second, the program must be structured to provide student teachers with the knowledge and skills necessary not only to carry out the underlying performance task, but also to document that performance in ways amenable for valid and reliable scoring. Third, the program must be structured to allow practice and feedback on these sets of knowledge and skills prior to a final independent effort. Fourth, scoring rubrics must be developed that, when used as designed, produce valid scores on the domains of the teaching task being assessed. Finally, a training program must be developed and implemented to ensure scores are reliable enough to use within a high stakes licensure decision context.

Within this context, faculty from the School of Education and Teaching Research spent the next three years working to create these conditions. The following timeline summarizes TEP-I work.

Phase I: Building the foundation - reacting to fall, 1994 National Review of TWSM (September 1995 - September 1996)

- A list of questions reviewers felt were yet to be addressed around TWSM was summarized by TEP Staff (see Schalock & Schalock, 1995).
- A work group of TEP staff and School of Education faculty was convened to address these questions and propose definitions, components and criteria for TWSM to satisfy these questions.
- Drafts of responses and rationales were developed by work group.
- Drafts were reviewed and refinements suggested through a series of whole faculty seminars over the course of the year that gained both individual, divisional, cross divisional and entire faculty input and consensus.
- A document was developed capturing these agreements (see Girod, Ayres, Ling, McConney, Schalock, & Wright, 1996).
- The agreements represented within this document were translated into changes in the Student Teacher Handbook, thus giving students developing work samples a clearer picture of expectations.

Phase II. Refinements in TWSM measures to reflect newly agreed upon definitions, components and criteria, as well as minimal variance noted in existing measures. (June 1996 - June 1997)

- Beginning drafts of a new set of measures more in line with processes, definitions, components and criteria associated with TWSM.
- Pilot use of measures by TEP staff on one work sample areas identified where disagreement would be likely by scorers. Beginning development of training session for scorers requirements.
- Development of a document bringing together processed, definitions, components, criteria, measures and uses underlying TWSM (see Ayres, Girod, McConney, Schalock, Schalock, & Wright, 1996).

- Drafts of measures were used to assess work samples by faculty to facilitate the review and refinement by the School of Education faculty through a series of whole faculty seminars over the course of the year that gained both individual, divisional, cross divisional and entire faculty input and consensus.
- Draft measures refined to reflect faculty input.
- Focus group sessions conducted with recent graduates around experience with and impact of TWSM, among other things.

Phase III: Redesign of the School of Education Program(s) to reflect new requirements by state licensing agency (TSPC) (January 1997 - August 1998)

- Considerable review of literature on teacher preparation, programs, teacher roles and functions, etc. conducted to guide the redesign effort.
- Initial cross-match between redesign and existing TWSM definitions, components, criteria, and measures conducted.
- Assistance in developing an overall framework for assessing the progress of intending teachers through the redesigned “proficiency-based” program.
- Work with OACTE in developing common frame of reference around TWSM across all teacher preparation programs in Oregon, both private and public.
- On-going support of new program implementation and use and refinement of overall assessment framework.
- Assistance in successful NCATE and TSPC re-accreditation efforts.

Phase IV: Completing the validation process (June, 1998 -)

- Submission of proposal for research study on impact of TWSM (TEP-II)
- Revisiting the alignment of TWSM processes, products and measures with redesigned programs
- Revisiting the reliability of scores from TWSM measures

In the end, the National Advisory Panel deemed the methodology as sufficiently robust, valid and reliable to be used as one measure that could be used to justify a high stake decision around licensure, if all the conditions were met. They were clear, however, that in and of itself, it could not be used to make licensure decisions, but that it could and should play a prominent role in a multiple measure approach to licensure decision.

TEP-II (1999-2003)

The work of TEP-I went far beyond TWSM. A developmental assessment system was developed to monitor progress of prospective teachers across all areas of knowledge and skill. Also, during this time the teacher preparation program was fundamentally redesigned to become an outcomes and standards-based program. The funder of TEP-I was interested in knowing what difference such a program design made in the thinking and practice of beginning teachers and the learning of their students. TEP-II was a longitudinal research project designed to answer that question, within the parameters of elementary schools in Oregon. While TWSM was one component of such a program design, it was not the only component. TEP-II, then changed the

focus of research on TWSM to research using TWSM as a major data collection tool for research on the effects of teacher preparation.

TEP-II was a relatively small sample size investigation designed to determine whether teacher preparation programs having specific policy related characteristics had any traceable and lasting effects on the thinking and practice of beginning teachers and the learning of their students. The three characteristics of teacher preparation programs of interest were: A) A strong alignment with Oregon's design for standards-based schools; B) An overt focus on connecting teaching and learning, especially through teacher work sampling; and C) An on-going developmental assessment of candidates against clear and public performance standards resembling those encountered in Oregon's K-12 schools.

Within this overall framework we were primarily interested in whether differently prepared teachers thought about and practiced teaching differently as first year teachers, and whether there was a corresponding difference in the level and quality of their students' learning. We were also interested in determining whether any first year effects were maintained in subsequent years of teaching, and if not, to determine those factors contributing to this result. From a theoretical and empirical standpoint we had some confidence that exposure to these program characteristics was a departure from traditional preparation, and was likely to result in teachers who thought about and practiced teaching differently and in ways that would support student learning in a standards-based school.

To answer these questions, a longitudinal, casual-comparative design was employed to follow 76 beginning elementary teachers through their first three years of teaching. Participating teachers experienced a range of preparation programs, both within Oregon and in other states. Through a combination of expert third party ratings, extensive content analyses of program materials and interviews with participating subjects, participant programs were sorted by level of exposure to the three program emphases of interest. These analyses and resulting ratings resulted in 13 of the 76 first year teachers being prepared in programs providing low exposure; 20 in programs providing moderate exposure; 16 high exposure; and 27 very high exposure. The study included 60-second year teachers. Due to attrition, from both the demands of the study and the profession, we were able to follow an intact group of 45 teachers across all three years.

To adequately capture the full range of variables related to the thinking and practice of beginning teachers, the learning of their students and the related affective, contextual and developmental variables of interest, we purposefully employed a mixed-method approach to data collection.

During the participants' first year of teaching we conducted extensive structured interviews; administered a suite of attitudinal surveys; and, administered a number of surveys and questionnaires to obtain descriptions of participants teaching contexts, professional development activities, and perceived level of competence. We also conducted focused classroom observations, and asked that a modified teacher work sample be prepared. These data collection activities were repeated for second year teachers participating in the study, with additional classroom observations and focus group interviews conducted in lieu of individual interviews. In the third year we continued to conduct multiple observations, but did not ask participants to prepare a third work sample.

The major conclusion reached was that the emphases in teacher preparation investigated were positively related to the classroom performance of first year elementary teachers, with effect sizes in the moderate to large range. These positive relationships, however, essentially disappeared in second year teachers and in some cases reversed themselves in third year teachers. No corresponding positive relationships between measures of student learning and program emphases or observed classroom performance were found. In part we believe this was due to limitations in obtaining consistent, high quality measures of student learning across participants. Further, there were a number of unanticipated findings related to the strength of contextual and affective variables on the development of teachers (Schalock & Schalock, 2004).

In using TWSM in the context of conducting this research on early career teachers it became obvious that many of the conditions necessary for it being a valid measure within the context of teacher preparation do not necessarily hold with practicing teachers. Nor does the level of control and standardization that can be exercised within a teacher preparation program exist in other than a randomized control trial of a specific teaching/teacher intervention. This was problematic for several reasons.

- First, the protocol used to guide the development of the work sample could not be prescriptive. Within the context of a teacher preparation program, prospective teachers can be asked to perform specific tasks and the methodology can be developed around those specific performance tasks. Within the context of the TEP-II research we wished to determine what beginning teachers actually did and how well they did it, not how well they could carry out a pre-determined set of performance tasks. Thus, the actual contents of the work samples differed to reflect the idiosyncratic nature of each teacher's practice.
- Because we could not fully define the performance tasks we could not specify scoring rubrics as fully as can be done in the context of a teacher preparation program. While we could have confidence in the reliability of the scores generated through the TWSM rubrics, we could have less confidence in their validity. Further research is needed to fully determine the conditions necessary for TWSM generating valid and reliable scores within the context of descriptive research. Certainly, the conditions necessary for its valid use in licensure decisions do not hold in field research.
- We do believe, however, that the methodology can be used within the context of research on specific interventions (as, for example, defined within the requirements of the Institute for Education Sciences).
- Finally, the burden placed on teachers, especially beginning teachers, to complete a quality work sample is simply an impossible task for many. There are no high stakes attached to completing a work sample, or at least a quality work sample, and a hefty stipend for completing did not provide enough incentive for a number of participants.

We did find, however, high levels of agreement around what beginning teachers said they did (as measured through TWSM) and what they were observed doing in the classroom (as measured through third party structured observations). Thus, even though there are a number of problems inherent in using TWSM in the context of research on practicing teachers, the methodology did, in this case, generate data consistent with that generated by trained observers.

(2003- and beyond)

Through several high profile efforts such as the AERA Consensus Panel on Teacher Education and the National Academy of Education Committee on Teacher Education, the field is poised like never before to move forward on the theoretical and empirical linkages between teacher preparation, practice, and P-12 student learning. It is well-recognized that few comprehensive data warehouses exist that will allow the kind of articulation necessary to make these connections (see the work of the Center for Teaching Quality and Wineburg, 2006). The Ohio Teacher Quality Partnership (Lasley, Siedentop, and Yinger, 2006) will emerge as an important contribution in this regard as it connects teacher preparation institutions, the Ohio Department of Education, and other state education policy makers in ways never before aligned. An incredible amount of time and dollars undergird these efforts and the field of teacher education is anxious for these efforts to bear fruit.

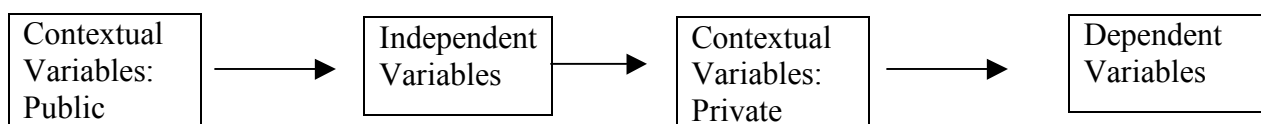
A small number of teacher preparation programs have been able to chain the long list of sophisticated variables to answer important questions about teacher preparation but, as a field, there is much definitional and validation work to be done. Teacher work sampling remains one of the few tools to assist in this process. If we can characterize the work described to this point as mostly establishing methodological consensus and empirical validation, the work described next can be characterized as most directly linked to uses in teacher preparation to demonstrate increased student achievement.

Exploring the use of a classroom simulation

The *Cook School District* simulation, designed at Western Oregon University, is a practice device for teacher education candidates acquiring skills in designing and implementing a teacher work sample (TWS) within their initial licensure program. Most faculty in teacher education know well the complexities of the knowledge and skills they expect their candidates to learn. Many faculty wish they had better ways to provide opportunities for their students to practice and receive feedback prior to stepping into a classroom. But tightly organized practice opportunities for teacher education students, prior to practica, are difficult to find.

Cook School District is a web-based, interactive system allowing students to practice the skills necessary in the design and implementation of a TWS. The basic intent of a TWS is to encourage the drawing of connections between P-12 student achievement and teacher performance. The operation of the simulation is controlled by three sets of variables. Two boxes within Figure 1 portray only one variable; i.e., boxes 1 and 3 both display “Contextual variables” and are thought of as a single variable.

Figure 1: Three variable sets within the algorithm for generating achievement and on-task responses



The independent variables include six elements candidate users manipulate as they design their instruction. The users' descriptions of each variable become part of the interactions within the algorithm that are utilized to develop behavior patterns for the simulated students. The specific components of the six independent variables are shown in Table 1. Student users select from the specific elements list descriptions of their decisions regarding the independent variables.

Table 1: Independent variables within the achievement algorithm

Independent Variables	Specific Elements
Test sequence	Pretest, formative, posttest
Item type	Oral, written, performance, constructed
Curriculum area	Math, science, art, health, wood shop, etc.
Domain and level	Attitude--general and self-confidence; Cognitive--remember, understand apply, analyze, evaluate, create, and; Physical--simple and complex
Instructional strategy	Assessing, computer-assisted instruction, discussing, feedback and recognition, lecturing/explaining, modeling, problem solving, seatwork, setting the stage, wrapping up
Homework	Amount, if any, the teacher expects of the students

The contextual variables include two types—public and private. The public variables are those where users read the simulated students' cumulative files. The contextual variables include information such as:

- Parents'/guardians' names, addresses, and occupations;
- Siblings—number of siblings and student's birth order;
- Health and demeanor reports from previous years;
- Academic records such as attendance, grades from previous years, standardized test performance, current class schedule, special program inclusion, if any, and teacher comments; and,
- Activities—athletic, special talent activities, and membership in leadership/service groups sponsored by the school.

The public variables are presented to give the teacher candidate user an overview of factors that might be expected to influence how students would perform in class.

The private files, box 3 in Fig. 1 above, are those which bear directly on a simulated student's academic and on-task performance. Each simulated student is based on a real person and was described by an author who knew the person well. The simulated students were described in terms of how they would respond to each of the specific elements which correspond to the six independent variables—see Table 1 above. The authors were asked to state the level of success the simulated student would likely have when each of the specific elements was encountered by the simulated student. For example, authors were asked, for Item Type, the score range a specific pupil would have to each of the assessment choices; i.e., if a shy child was being described, the achievement score range would likely be low.

Dependent variables within the *Cook School District* simulation are of two kinds—achievement scores and on-task responses. Both scores are influenced by the candidate’s use of the independent variables. The idea is, of course, to manipulate the independent variables, in ways informed by an understanding of the public contextual variables, to influence the dependent variables. The private contextual variables make up the “engine” of the simulation – the algorithmic model for determining achievement and on-task responses.

Initial pilot testing of the *Cook School District* simulation suggest that, when compared against a control group of non-simulation users, users of the *Cook* simulation:

- Showed statistically significantly higher ratings of real-world teaching evaluations particularly in the areas of planning, establishing a classroom climate conducive to learning, and ability to evaluate pupil learning.
- Pursued statistically significantly greater portion of higher order learning outcomes and helped P-12 students meet them with greater regularity.
- Were able to produce real teacher work samples that were rated statistically significantly better than treatment group students.
- Showed statistically significantly higher ratings of perceived skillfulness in connecting teaching and learning as well as ratings of the value held for the skills necessary to connect teaching and learning.
- Could analyze and evaluate with more sophistication simulated teaching scenarios and more consistently labeled planning, individualizing for learners, and evaluation as critical processes in successful teaching and learning.

Each of these findings should be considered tentative and in need of cross validation but they are exciting in their ability to connect preparation effects, through to teaching practices, and on to student learning.

Oregon Collaborative Research Initiative (OCRI)

Given the developmental and initial empirical work of the TEP studies, it became necessary to begin to think about how to generate data across programs, cross-validating the TEP research in multiple contexts. In recognition of this, the two state-level affiliates of ATE and AACTE helped design and launch the Oregon Collaborative Research Initiative (OCRI). The leadership within these groups, representing all 16 teacher preparation programs in the state, hired a research design to assist in addressing essential linkages from preparation, to practice, through to P-12 student learning. What emerged was organization around three sets of questions or issues. Each are discussed below:

First, it is essential to understand the range of perceptions and experiences of both teacher candidates and teacher educators as they employ the processes of teacher work sampling. Though the state licensing commission in Oregon requires that work samples be prepared by candidates seeking licensure, the precise nature of this work sample and how it is described, articulated, and evaluated is left to individual teacher preparation programs. This presents an obvious challenge to those interested in looking across programs to speak broadly about the

efficacy of the methodology so understanding the range of experiences is critical. Research questions in this area include:

- Q1.** How do candidates view the work sample?
 - Q1a.** Do candidates view the work sample as *authentic* (i.e., meaningful, worth their time, useful, and having applicability to their professional preparation and work as teachers)?
 - Q1b.** Do these perceptions change as students move through the program and enter the profession (e.g., How might candidates initial experiences compare with later experiences in the program and then one, two, three years out of the program)?
 - Q1c.** What aspects of work sampling do candidates see as most valuable?
 - Q1d.** What concerns do candidates have about work sampling and its use in teacher preparation and licensure?
- Q2.** How do teacher educators view work sampling?
 - Q2a.** How central is the role of work sampling in teacher preparation programs?
 - Q2b.** How integrated is work sampling in teacher preparation programs?
 - Q2c.** How is information used from work samples to inform coursework and field experiences?
 - Q2d.** What aspects of work sampling seem to work well and not so well in supporting teacher development?
 - Q2e.** What questions or concerns do faculty have about work sampling?

Second, as with any innovation work sampling has been adapted, extended, and refined in ways that support the intellectual character, values, and demands of varying institutions and professional contexts. Specifically, work sample requirements, prompts, and evaluative tools have been modified for use in addressing issues of cultural competence, for use in various special education settings, and in specific content areas like mathematics and science. A rich understanding of the role of work sampling in Oregon teacher preparation and licensure requires an analysis of these elaborations. Research questions in this strand include:

- Q1.** How has the work sample been elaborated across various teacher education institutions?
 - Q1a.** What specifically has been added to the work sample?
 - Q1b.** What prompted this particular elaboration of the work sample?
- Q2.** What, if any, similarities are there to the elaborations across institutions?
- Q3.** How successful have these elaborations been in terms of effectively addressing the demands of the unique context?

Third, it is essential for Oregon to add what is known from our long-term commitment to the processes of work sampling to the professional knowledge emerging around the linkage between teacher preparation, teaching practices, and P-12 student learning. Value-added modeling has helped significantly link teaching to student learning but provides us little information around teacher preparation effects. In recognition of this, a third strand on the impact of work sampling on teacher practice and P-12 student learning was also designed. Research questions from this strand include:

- Q1.** What is the impact on K-12 Students?
- Q1a.** What techniques do candidates use to assess the impact of their instruction? And how sound are those techniques?
- Q1b.** How do candidates disaggregate data and use that data for making inferences about learning?
- Q1c.** How well do candidates support student learning claims and inferences?
- Q1d.** How do social factors (e.g., characteristics of the learning context) impact candidates' ability to represent and interpret student learning data?
- Q1e.** How do candidates' personal factors (e.g., assessment-efficacy) impact their ability to represent and interpret student learning data?
- Q2.** What is the impact on teacher candidates?
- Q2a.** How does TWS preparation influence candidates approach to student learning assessment and interpretation over time?
- Q2b.** What, if any, instructional behaviors of post-graduates can be attributed to TWS?
- Q3c.** How might candidates' performance on TWS predict other outcomes of teacher development?

Together, through the Oregon Collaborative Research Initiative, we hope to answer several critical questions related to the use and value of teacher work sampling in teacher preparation.

Efforts moving beyond Oregon

Most recently organized, and with the largest project scope, is the Coalition for Connecting Teaching, Preparation, and Learning (<http://www.cctpl.org>). This group includes researchers, educators, and policy-makers from across the country committed to demonstrating the efficacy of teacher preparation by chaining the necessary inferences to drill down from preparation through to student learning. The work of the coalition has, to this point, centered on defining the contexts in which teaching and learning in a standards-based system occur and the conditions necessary (focusing first on teacher-level variables) for learning to occur. The group has pushed for conceptual clarity around ideas that incorporate a developmental perspective on teaching and the career of teachers, that considers seriously the enabling knowledge and skills necessary for good teaching and sophisticated P-12 student learning, as well as the demands and challenges of organizing to build theory and gather empirical support for a clinical teaching profession. It is the ultimate goal of the coalition to strengthen the body of empirically driven theory in teacher education and to transform the field of teacher education from one driven by hunches and well-meaning individuals to one driven by a scientific vision of teacher preparation and individuals who understand and are committed to an internally consistent set of concepts, standards, and measures. It is our hope that a series of outcomes for the work of this group could include but are not limited to (there's much more to this list):

- Operational definitions of variables and their relationships in teaching and learning within standards-based systems of teaching and learning
- An associated index of key terms, concepts, and their conceptual and empirical relationships
- A mapping of the relationships between theoretical constructs, measures, programs, policies and related research that meaningfully connect teaching, preparation and P-12 learning

- A catalog of empirically validated measures of these variables and standards for their appropriate use

It is possible to consider the outcome of these efforts as the design and validation of a series of clinical teaching tutorials that will define, demonstrate, and support the preparation of teachers for standards-based schools who will be able to ensure learning in P-12 students. The work of the CCTPL group has just begun – more as it unfolds.

Several other critical efforts have been launched to build the kinds of systems necessary, to do the kind of work required, to effectively connect (theoretically and empirically) teacher preparation, teaching practices, and P-12 student learning. Many pages could be dedicated to the sophisticated and well constructed efforts of such organizations and collaborations as the Carnegie Foundation's Delta Project led by Lee Shulman, the Carnegie Corporation's Teachers for a New Era and the Learning Network project, the Ohio Teacher Quality Partnership, and the work of the Oklahoma Commission for Teacher Preparation just to name a few. Each of these efforts share a common commitment to an analysis of teacher preparation through to P-12 student learning – though with varying methodological and pedagogical commitments. It is truly an exciting time for teacher education research and theory development and teacher work sampling continues to play an important and increasingly high profile role in this regard.

Summary and conclusions

It is obvious that teacher work sampling has provided a rich venue for thinking carefully about the role of teachers and teaching in P-12 student learning and has helped to move teacher preparation from a disparate set of classes and disconnected field experiences to a more tightly designed system, focused on the goal of connecting teaching and learning and standards-based schools. TWSM is also providing cautious confidence in the ability to prepare teachers who make a difference in the learning of children. TWSM has been empirically validated and a system of scoring and prompts have been developed to help meet appropriate conditions for use. A companion observation system and, in some cases, a companion data management system have been developed to afford institutional effectiveness data in real time (see EDSMART at Western Oregon University). Smaller scale projects have continued to drill down from these larger efforts into explorations of individual preparation experiences and their impact on practice and P-12 learning. Other universities and states have adapted TWSM and have extended the validation and conditions for use in important ways. It is the hope of the Coalition for Connecting Teaching, Preparation, and Learning (CCTPL), through their collaboration with individuals and institutions with an existing commitment to connecting teaching and learning through work sample methodologies, to begin systematically moving forward on the development of theoretical and empirical connections necessary to “close the loop.” Significant progress has been made in the last 20 years but the field of teacher education and our methodologies must continue to be refined and extended to meet the ethical obligation of ensuring that all children learn to their highest potential.

References

- Ayres, R., Girod, G., McConney, A., Schalock, M.D., Schalock, D., & Wright, D. (1996). A Guide to Teacher Work Sample Methodology. Monmouth, OR: Teaching Research Division and School of Education, Western Oregon State College.
- Berliner, D. (2002). Educational research: The hardest science of all. Educational Researcher, 31(8), 18-20.
- Cochran-Smith, M. (2005). Studying teacher education: What we know and need to know. Journal of Teacher Education, 56(4), 301-306.
- Darling-Hammond, L. (1997). The right to learn. San Francisco: Jossey-Bass.
- Elliott, E. (2004). "Student learning in NCATE accreditation". Quality Teaching. Washington, DC: National Council for Accreditation of Teacher Education 13(1), p. 4-5.
- Fetler M. (1999). High school staff characteristics and mathematics test results. Education Policy Analysis Archives, 7(9): <http://epaa.asu.edu/epaa/v7n9.html>
- Girod, G. (Ed.) (2002). Connecting teaching and learning. Washington, D.C.: American Association of Colleges for Teacher Education.
- Girod, G., Ayres, R., Ling, M., McConney, A., Schalock, M.D., and Wright, D. (1996). The use of work samples at Western Oregon State College. Monmouth, OR: Teacher Effectiveness Project, Center for Teaching and Learning, Western Oregon State College.
- Lasley, T., Siedentop, D., & Yinger, R. (2006). A systemic approach to enhancing teaching quality: The Ohio model. Journal of Teacher Education, 57(1), 13-21.
- Sanders, W., and Horn, S. (1994). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. Journal of Personnel Evaluation in Education, 12 (3), 247-256.
- Schalock, M. (1998). "Accountability, student learning, and the preparation and licensure of teachers: Oregon's teacher work sample methodology". Journal of Personnel Evaluation in Education, 12:3, 269-285.
- Schalock, H.D., and Schalock, M.D. (1995). Proceedings of a Conference: Advances in Theory and Research on Teacher Effectiveness. Monmouth, OR: Western Oregon State College.)
- Schalock, M.D., and Schalock, H.D. (2004). The longitudinal effects of teacher preparation on the practices and beliefs of new teachers and the learning of their students: Final project report [On-line]. Available at: <http://www.tr.wou.edu/tep/products.html>
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. Education Policy Analysis Archives, 10(12); retrieved from <http://epaa.asu.edu/epaa/v10n12/>
- Wineburg, M. (2006). Evidence in teacher preparation. Establishing a framework for accountability. Journal of Teacher Education, 57(1), 51-64.