

Teacher Work Sample Methodology: Displaying Accountability Of U.S. Teacher Education Program Effectiveness

Donna Cooner, Colorado State University, USA
Cerissa Stevenson, Colorado State University, USA
Heidi Frederiksen, Colorado State University, USA

ABSTRACT

Productive strategies for evaluating outcomes in teacher licensure programs are becoming increasingly important in the education field. Research data from 492 teacher licensure program completers from 2006-2009 in the School of Education, Colorado State University, Fort Collins, Colorado (USA) was compiled after accreditation by the Colorado Department of Higher Education (CDHE) and the Teacher Education Accreditation Council (TEAC). Three categories including subject matter knowledge, pedagogy, and caring teaching skills were acknowledged in accreditation. An important pedagogy experience for students during the licensure program was teacher work sample methodology. This article discusses teacher work sample methodology, explains the experience of building a teacher work sample (TWS) and presents data regarding students' results from TWS rubrics. Findings demonstrate teacher education program effectiveness and highlight the effective use of TWSs to meet accreditation requirements.

Keywords: Teacher Work Sample Methodology; Teacher Work Sample; Teacher Education; Accountability; Teacher Preparation Program Effectiveness

INTRODUCTION

University-based teacher education programs in the past decade have been called on repeatedly to prove their effectiveness, from state legislators to alternative education program entrepreneurs to parents (CO RttT proposal, 2010; Educational Testing Service, 2002; Feistritzer, 2004; Hess, 2001; Paige, 2002). The U.S. Secretary of Education, Arne Duncan (2009) outlined the urgency to improve teacher preparation in a speech delivered at Teachers College, Columbia University:

To keep America competitive, and to make the American dream of equal educational opportunity a reality, we need to recruit, reward, train, learn from, and honor a new generation of talented teachers. But the bar must be raised for successful teacher preparation programs because we ask much more of teachers today than even a decade ago. Today teachers are asked to achieve significant academic growth for all students at the same time that they instruct students with ever-more diverse needs. Teaching has never been more difficult, it has never been more important, and the desperate need for more student success has never been so urgent. Are we adequately preparing future teachers to win this critical battle?

Reacting to this increased pressure, university administrators and teacher educators are working hard to build accountability and prove effectiveness (Wineburg, 2006). Teacher education institutions nationwide are shifting from “input to output-based programs [and are] trying to define what it means to be a ‘value-added’ endeavor...linked to student learning and school change” (Cochran-Smith, 2006, p.2).

Accreditation agencies, associations, legislators and experts in the field set standards for excellence and identify benchmarks that indicate teacher candidate quality. Three outcomes are consistently noted as indicators of

teacher candidate success: 1) content knowledge; 2) pedagogy that integrates theory and practice and meets the needs of diverse learners; and 3) caring, emergent professionalism. The National Council for Accreditation of Teacher Education (NCATE) is requiring that colleges of education reinvent themselves to prepare candidates with high performance standards, and document candidates' teaching ability through clear assessments including performance-oriented results (Wise & Leibbrand, 2001). NCATE's Standard 1, *Candidate Knowledge, Skills, and Professional Dispositions* states: Candidates preparing to work in schools as teachers or other school professionals know and demonstrate the content knowledge, pedagogical content knowledge and skills, pedagogical and professional knowledge and skills, and professional dispositions necessary to help all students learn (NCATE, 2008). Other accrediting agencies, like the Teacher Education Accreditation Council (TEAC) have adopted similar standards. They recognize that teachers need subject matter knowledge (quality principle 1.1), pedagogical knowledge (quality principle 1.2), and caring teaching skills (quality principle 1.3) (TEAC, 2007). *The Teacher Educator's Handbook* (Murray, 1996) has chapters on subject matter knowledge, pedagogy, and meeting the needs of diverse learners. Darling-Hammond's (2005) *Framework for Understanding Teaching and Learning* includes "Knowledge of subject matter and curriculum goals; Knowledge of teaching; and Knowledge of learners and their development in social context" (p. 5).

TEAC asserts that program candidates must understand the subject matter they will teach. Their first quality principle is *Subject Matter Knowledge*. Subject matter knowledge has increased in importance because the standards for learning are higher than they ever have been before (Darling-Hammond, 2005). It has also been argued that understanding teacher content knowledge is essential for secondary school reform (Grossman & Stodolsky, 1995).

Pedagogical practices are also imperative to consider when evaluating teacher candidate quality. The American Association of Colleges for Teacher Education (AACTE) is leading reform efforts that "go beyond content knowledge" and prepare teachers to differentiate their instruction to reach all children, especially those most at risk for school failure: children with disabilities, English language learners, and children from low-income homes. This is also reflected in TEAC's second quality principle: *Pedagogy*. TEAC requires that "program candidates must be able to convert their knowledge of subject matter into compelling lessons that meet the needs of a wide range of pupils and students" (2007). Darling-Hammond (2006) also highlights the increased demands on teachers to teach all students for higher order thinking and performance skills that formerly were only offered to a select group of students.

The third quality principle for TEAC (2007) is *Caring teaching skills*. The principle reads as follows:

Above all, teachers are expected to act on their knowledge in a caring and professional manner that would lead to appropriate levels of achievement for all their pupils.

Caring is a particular kind of relationship between the teacher and the student that is defined by the teacher's unconditional acceptance of the student, the teacher's intention to address the student's educational needs, the teacher's competence to meet those needs, and the student's recognition that the teacher cares.

Darling-Hammond, (2006) states that "programs must help teachers to develop the disposition to continue to seek answers to difficult problems of teaching and learning and the skills to learn from practice" (p. 5).

The driving question now is "How will we know when (and if) teachers know and can do what they ought to know and be able to do" (Cochran-Smith, 2006, p. 13). Literature suggests that the implementation of embedded assessments such as teacher work samples can contribute to programmatic changes necessary to meet accreditation standards (Youngs & Bird, 2010, p.187). Carefully designed assessments embedded within teacher education programs can promote pedagogical reasoning if they require candidates to gather information about their students' learning, formulate hypotheses to explain trends, and use those hypotheses to deliberately modify instruction and assessment (Youngs & Bird, 2010, p. 187).

Though measures such as portfolios and embedded assessments are more commonly being used as evaluations in teacher education, Youngs and Bird (2010) attest that there is little evidence that these assessments of

pre-service teachers are valid and reliable and are effective measures of instructional practice required by accrediting agencies. This investigation was conducted to evaluate the validity and reliability of one university's measure of instructional practice and to examine evidence outcomes of a teacher preparation program's claim that graduates are competent in subject matter knowledge, pedagogy that integrates theory and practice designed to meet the needs of diverse learners, and caring, emergent professional. Teacher Work Sample (TWS) Methodology was utilized. TWSs have been found to be one of four primary methods for establishing teacher education program effectiveness (Wineburg, 2006). "While many programs demonstrate teacher candidates' impact on student learning by using pretests and posttests of students, TWS represents something more...it considers the context, planning, student learning, and reflection" (Lorson, 2010, p. 55).

COLORADO STATE UNIVERSITY'S PROGRAM

The context of this study is a land-grant and space-grant (partially funded for space-related research), Carnegie Doctoral/Research University-Extension. There are 20,720 undergraduate students, 3,690 graduate students, and 537 professional veterinary medicine students. The Teacher Licensure Program is one of eight academic units in the College of Applied Human Science and prepares an average (from 2000-2009) of 191 educators a year. The program is aligned with state standards and is affiliated with the National Network of Educational Renewal.

The Colorado State University (CSU) teacher education program is designed as a developmental progression of coursework and field experiences that leads candidates through the initial process of learning to teach and culminates in the final recommendation for teacher licensure. The program is delivered in four phases of study reinforced throughout by a philosophical and programmatic core of learning based on the following three concepts: national (USA), state and institutional standards; partnerships between the university and local school communities; and experiential learning opportunities for teacher candidates.

Faculty members in teacher education programs have had a long-term responsibility to future teachers to not only transmit information, but to model what their candidates are expected to do (Goodlad, Soder, & Sirotnik, 1990). The professional education faculty at CSU believes that creating an invitational environment and translating critical theory to classroom practice creates high competencies. Through experiencing these practices and environments, teacher licensure students are believed to better comprehend the role of the teacher as a facilitator of student success.

Goodlad's (1990) concept of effective modeling is similarly addressed by Vygotsky's (1962) description of relational imitation and John Dewey's (1938) notion of learning through direct experiences. Teacher candidates at CSU complete at least 800 hours of field experience. In their endeavor to identify specific instructional features promoting meaningful growth in teacher education candidates, Jensen and Winitzky (1999) examined over 43 studies on educational improvement. Thirty-two of these investigations reported meaningful learning in candidates with teacher training programs that emphasized course content used in context, repeated reflection, and modeling by faculty and other professional educators. As Goodlad (1990) surmised, "We recommend, then, that the responsible faculty plan not just a sequence of courses and field experiences but deliberate demonstration of pedagogical procedures their students will be expected to use in the practice part of their preparation programs" (p. 291). Dual pedagogy narrows the theory to practice divide that some students believe exists between their teacher education programs and their professional teaching experiences (Ketter & Stoffel, 2008).

CSU's program is affiliated with the National Network of Educational Renewal. In the mid-1980s, thanks to the work of organizations such as the Holmes Group (now the Holmes Partnership) and the National Network for Educational Renewal, the term "Professional Development School" (PDS) emerged as a part of the nation's educational discourse. Unique and particularly intense school-university collaborations, PDSs were designed to accomplish a four-fold agenda: preparing future educators, providing current educators with ongoing professional development, encouraging joint school-university faculty investigation of education-related issues, and promoting the learning of PK-12 students. In addition to identifying the distinctive nature of the PDS relationship, the mission should also focus on two overarching goals: the advancement of the education profession and the improvement of P-12 learning. In furthering the education profession, the PDS relationship should be all-inclusive in its promotion

of professional growth across the continuum of pre-service teacher candidates, in-service educators, and college/university faculty and administrators.

Colorado State University's (CSU) Teacher Licensure Program directly aligns its practice with the Colorado Department of Education Performance-Based Standards for Colorado Teachers (2000) and has a firm commitment to engaged, standards-based, and inquiry-based teacher education. The four-phase program was implemented to meet the principles of the Performance-Based Standards for Colorado Teachers (PBSCT):

- (1) *Standard One: Knowledge of Literacy.* The teacher shall be knowledgeable about student literacy development in reading, writing, speaking, viewing, and listening.
- (2) *Standard Two: Knowledge of Mathematics.* The teacher shall be knowledgeable about mathematics and mathematics instruction.
- (3) *Standard Three: Knowledge of Standards and Assessment.* The teacher shall be knowledgeable about strategies, planning practices, assessment techniques, and appropriate accommodations to ensure student learning in a standards-based curriculum.
- (4) *Standard Four: Knowledge of Content.* The elementary teacher is knowledgeable, in addition to literacy and mathematics in the following content areas: civics, economics, foreign language, geography, history, science, music, visual arts, and physical education. Middle school and secondary content teachers shall be knowledgeable in literacy and mathematics and expert in their content endorsement area(s).
- (5) *Standard Five: Knowledge of Classroom and Instructional Management.* The teacher is knowledgeable about classroom practice in order to successfully manage time, communications, and record keeping procedures that will support and enhance student learning.
- (6) *Standard Six: Knowledge of Individualization of Instruction.* The teacher is responsive to the needs and experiences children bring to the classroom, including those based on culture, community, ethnicity, economics, linguistics, and innate learning abilities. The teacher is knowledgeable about learning exceptionalities and conditions that affect the rate and extent of student learning, and is able to adapt instruction for all learners.
- (7) *Standard Seven: Knowledge of Technology.* The teacher is skilled in technology and is knowledgeable about using technology to support instruction and enhance student learning.
- (8) *Standard Eight: Democracy, Educational Governance and Careers in Teaching.*
The teacher recognizes the school's role in teaching and perpetuating our democratic system. The teacher knows the relationships among the various governmental entities that create laws, rules, regulations, and policies that determine educational practices.

Research data from 492 teacher licensure pre-service students in the CSU School of Education from 2006-2009 was compiled after the Colorado Department of Higher Education (CDHE) program evaluation in 2008 and for the Teacher Education Accreditation Council (TEAC) process in 2010. Three categories including subject matter knowledge, pedagogy, and caring teaching skills were acknowledged in accreditation. Often there are more outcomes in subject matter knowledge than pedagogy and caring teaching skills while students progress through their teacher licensure program. Teacher Work Sample data significantly contributed to the overall evaluation of the program and increased the understanding of teacher candidate development.

Teacher Work Samples

An important pedagogy experience for students during the licensure program is teacher work sample methodology (Darling-Hammond, 2006; McConney, Schalock & Schalock, 1998). Teacher work sample methodology emphasizes attention to pupil learning for beginning teachers, is designed to meet accreditation requirements and standards, and gives evidence for its reliability and validity in evaluating the performance of beginning teachers (Denner, Norman, Salzman, Pankratz, & Evans, 2004). Originally conceived at Western Oregon State University (Perry, Smith, & Woods McConney, 1998), the TWS provides a suitable strategy for measuring student success and then for the candidate, through reflection, to better understand processes for enhancing student learning in the future.

TWSs are modeled in courses during the initial phase of preparation in CSU's program. Then formative TWSs are developed by candidates in the next two phases. Candidates develop summative TWSs during the final phase; the student teaching experience. The summative student teaching TWS gives prospective teachers experiences in designating standards-linked learning outcomes to be accomplished through a two to five week unit of instruction; developing plans for instruction and assessment that are aligned with the outcomes desired; and collecting, interpreting, and reflecting on evidence of student progress toward outcome attainment.

The TWS allows prospective teachers to assess the success of their teaching, the progress of their students in relation to the objectives and standards delineated in the planning, and then provide them the ability to reflect and make the changes necessary for improved teaching and student success. The TWS gives teacher education candidates the opportunity to demonstrate that they are in fact able to foster learning gains in students (Denner, Salzman, & Harris, 2002). An effective TWS connects teaching and learning, enabling candidates to combine and blend an interdependent set of processes, products, results, and reflections.

Analyzing the process by which the TWS is developed and implemented, the reliability and validity of the assessment, and the impact on teacher candidate readiness is necessary to evaluate teacher preparation programs. The outcomes of these analyses inform teacher educators so they can implement meaningful pedagogical, programmatic and evaluative changes. With our focus of inquiry begin on teacher candidate development (aligned with TEAC principles), our research questions were:

- (1) Do teacher candidates understand the subject matter they intend on teaching?
- (2) Do teacher candidates utilize pedagogy that integrates theory and practice in efforts to meet the needs of diverse learners in real world educational contexts?
- (3) Are teacher licensure graduates caring, competent emergent professionals?

Two final research questions address the validity and reliability of the TWS:

- (4) Does the TWS truly measure the constructs mentioned in research questions 1 through 3?
- (5) Is there consistency within CSU's licensure program in assessing the development of teacher candidates?

METHODS

After completing formative TWSs during the initial phases, each student teacher produces a complete, summative Teacher Work Sample. The TWS is evaluated by a university supervisor and graded according to a shared rubric. This grade is 75% of the student's final grade for student teaching. The Teacher Work Sample serves as an "extended snapshot" of the student teaching experience. It allows prospective teachers to assess the success of their teaching, the progress of their students in relation to the objectives and standards delineated in the planning, and the ability to reflect and make the changes necessary for improved teaching and student success. A Teacher Work Sample connects teaching and learning, enabling students to combine and blend an interdependent set of processes, products, results, and reflections.

The objectives for the Teacher Work Sample at CSU are to specifically develop prospective teachers' abilities to:

- (1) learn to think reflectively about teaching—planning, instruction, assessment, management of the learning environment, communication, and professionalism.
- (2) integrate curriculum, instruction, and assessment.
- (3) employ decision-making processes that continuously search for the most appropriate fit among the various aspects of instructional context, subject matter knowledge, and repertoire of appropriate instructional strategies and procedures.
- (4) practice and obtain feedback about one's effectiveness as a teacher in fostering pupil progress in learning (formative) demonstrate and document effectiveness in fostering learning gains (summative).

As shown in Table 1, the teacher work sample scoring rubric has eight sections each scored separately for a total of 100 points.

Table 1. Teacher work sample sections and cut scores.

Section	Maximum Points	Advanced	Proficient	Developing
I – Setting and Context	10	10-9	8.5-8	7.5-0
II - Unit Topic and Rationale	5	5-4.5	4	3.5-0
III – Unit Standards, Goals, and Objectives	10	10-9	8.5-8	7.5-0
IV – Assessment Tools (Pre-, Post, and Instructional)	10	10-9	8.5-8	7.5-0
V– Instruction and Management Plan	30	30-27	26.5-24	23.5-0
VI – Assessment Data and Analysis	20	20-18	17.5-16	15.5-0
VII Student Teaching Experience Reflection	10	10-9	8.5-8	7.5-0
VIII – Organization and Presentation	5	5-4.5	4	3.5-0
Total	100			

The use of embedded assessments such as the TWS brings concern about the consistency with which multiple university supervisors can implement and score reliable and valid candidate assessments. The TWS is scored by several different faculty members each semester. To offset issues with reliability, training and discussion occurs each semester on scoring procedures and examples of student work from previous semesters discussed. Collaboration among university faculty can be seen as a property of teacher education programs that enable them to design and implement change (Younges & Bird, 2010, p. 191). Correlations among instructors' ratings of the components were considered for final evaluation of inter-rater reliability.

University supervisors, cooperating teachers and student teachers are provided with a copy of the evaluation forms prior to beginning student teaching. The rubric for scoring the TWS was developed by an expert faculty team and has been used to measure the components for the TWS for over five years. All performances are aligned with state standards and course outcomes, which increases content validity.

Participants

CSU's Teacher Licensure Program delivers 11 initial program options in secondary education, four K-12 program options in art, music, foreign language, and instructional technology (computing); and one early childhood education (ECE) program option at the undergraduate and post-bach level. Early childhood is not offered in the master's option. The Colorado legislature designated CSU as the only Colorado institution of higher education eligible to provide teacher licensing and credentialing programs for career and technical education (formerly known as vocational education). Program options at the undergraduate and post-bach level include agricultural education, business education, distributive (marketing) education, home economics (family and consumer sciences), industrial arts (trades and industry education), and technology education (engineering). Program completers from 2006-2009 were included for this inquiry. Table 2 includes the numbers of students per year in undergraduate, post-bachelors and graduate levels, as well as the number of students by content area. The graduate program was restructured in 2007, which is reflected in the small numbers of program completers in 2009. Subsequent years have reflected similar enrollment to those prior to 2007.

RESEARCH INSTRUMENT AND DATA GATHERING

Subject matter knowledge. As a part of the Teacher Work Sample, students must demonstrate their ability to plan lessons based on content standards and must demonstrate knowledge of content in their teaching sample. We examined the section titled *Element II-Unit Topic and Rational* of the TWS scoring rubric for this question (Maximum Score=5).

Pedagogy integrates theory and practice for diverse learners. Directly related to the student's ability to utilize pedagogy as a part of the Teacher Work Sample, students must demonstrate their ability in the areas of *III-Unit Standards, Goals and Objectives* (Maximum Score=10); *IV Assessment Tools* (Maximum Score=30 pts) and *V-Instruction and Management Plan* (Maximum score=30). All of these areas directly relate to the student's ability to integrate theory and practice in their utilization of pedagogy.

Caring, emergent professionals. As a part of the Teacher Work Sample, students must reflect on their teaching experiences about what went well and what they would do differently. Insightful, reflective teaching is an integral part of developing into a caring teacher. We examined the section titled *Element VII Student Teaching Experience Reflection* of the TWS scoring rubric for this question (Maximum Score=10).

Assessing Validity. Content validity ensures that the instrument used actually measures what it says it will. In this case, not only do the standards need to be linked to the TWS sections, but also the requirements for accreditation (knowledge, pedagogy and caring teaching skills) must link to the TWS.

The construct validity examines the meaning of the scores in terms of the intended Teacher Work Sample constructs. A factor analysis estimated the variance in each variable accounted for in each component. A high extraction number will indicate that the components represent the variable well. Correlations were also run between mean scores to analyze their relationships.

Assessing Reliability. To evaluate the consistency of scoring the TWS, comparisons among assessors are typically measured. A one-way ANOVA was run to compare scores between university supervisors.

Table 2 Admitted CSU Licensure Students in each Licensure Option

Licensure Option	Year of Completion				Total Number	Total %
	2006 N	2007 N	2008 N	2009 N		
Undergraduate	66	196	187	130	579	76%
Post-bach/licensure only	13	34	32	29	108	14%
Graduate	6	37	27	3	73	10%
Total N	85	267	246	162	760	100%

	2006	2007	2008	2009	Total
Agricultural Ed	6	10	9	5	30
Art	5	28	19	17	69
Family/Consumer	5	14	15	10	44
Early Childhood	2	26	29	22	79
English	19	46	45	28	138
Foreign Languages	1	15	11	9	36
Marketing/Business	1	11	5	3	20
Mathematics	3	17	25	9	54
Music	6	11	24	11	52
Science	9	39	29	27	104
Social Studies	26	37	30	32	125
Speech	2	3	1	0	6
Tech Ed	0	1	2	0	3
Total	85	267	246	162	760

FINDINGS

The descriptive statistics show that overall, students scored in the advanced area on the different work sample categories. An ANOVA indicated the *Topic and Rational* and *Instructional Plan* categories were significantly different (both were significant: $p = .02$) across the five TWS categories. Differences within each of the groups, however, were not statistically significant. Differences between years were not statistically significant from 2006-2008 ($p > .05$). Faculty changed rubrics in 2009 and scores for several categories were modified. Statistically this is verified with an ANOVA. Each TWS category that was changed was statistically significant: *Unit standards, goals and objectives* ($p = .002$); *Assessment tools* ($p = .006$); and *Instruction and management plan* ($p = .000$). Games-Howell Post-hoc test verifies that the difference was between 2006-2008 and 2009.

The following findings include the areas of the Teacher Work Sample that directly relate to the research questions one through three. Analysis of scores indicated that the means were in the advanced range for the majority of the TWS sections. Data were disaggregated by content area to identify any differences in student performance between the disciplines.

Evidence of students' understanding of the subject matter was analyzed using the scores from the *Unit Topic and Rationale* section of the TWS. This includes students' ability to: *clearly describe the topic taught; describe the relevance of the topic to the content standards and within the curricular sequence; and to explain the significance of the topic globally, personally, and specifically for this population.* Of the 492 program completers, the mean scores for all content areas were in the advanced range.

Evidence of teacher candidates' ability to utilize pedagogy that integrates theory and practice in efforts to meet the diverse learners in real world educational contexts included the *Unit Standards, Goals, and Objectives* section of the TWS which incorporates the following observed behaviors: *Goals and objectives are appropriate and aligned with State Content Standards; Goals and objectives are clearly stated; Literacy (speaking, writing, reading, or listening), numeracy, and technology are addressed by at least one goal and/or objective; Objectives address higher order thinking skills; Objectives allow students various means to demonstrate proficiency.* The mean scores for all content areas were in the advanced range.

The *Assessment Tools* section of the TWS included points awarded for the following observed behaviors: *Unit contains pre- and post-assessment and items are aligned with goals and objectives; Pre and post assessment items are either identical or parallel; a variety of assessment tools is utilized (e.g. traditional, authentic, self-assessment, and peer assessment); The expectations of alignment, clarity, and developmental appropriateness of all assessments are explained graphically and/or in essay form.* The mean score for all content areas was in the advanced range.

Points awarded for the *Instruction and Management Plan* section of the TWS incorporated the following observed behaviors; *Explain the sequence and organization of the unit; The unit must be cohesive rather than a collection of separate lessons; Demonstrate knowledge of content; Discuss methods for integrating literacy, numeracy, technology and other relevant subjects; Demonstrates at least three different instructional strategies, one of which is cooperative learning; Lesson plans must include all key components; Lesson plans are aligned with objectives; Lesson plans include method and criteria for assessment; Objectives and assessment are aligned; Lesson plans include strategies for management relevant to the lesson.* The mean score for all content areas were in the advanced range.

The evidence demonstrating that the teacher candidates are caring, competent emergent professionals included scores from the Teacher Work Sample section, *Student Teaching Experience Reflection*. In this section of the TWS the perspective teacher concretely reflects about the following issues: *What are the standards-linked learning outcomes I want my students to accomplish; What activities and instructional methodologies are appropriate or necessary for these students to achieve these outcomes; What resources and how much time do I need to implement these activities and methodologies; What assessment activities or methodologies are appropriate for these students and these outcomes when using these instructional methodologies; How successful was I at helping my students achieve the outcomes desired; What went right? What went wrong? Why?* Completers in all content areas combined, averaged greater than 90% for Experience and Reflection.

VALIDITY

Content Validity: Standards and accreditation requirements (knowledge, pedagogy, and caring teaching skills) were all linked to the TWS section, as show in Table 3. These findings provide strong evidence of the Teacher Work Sample's content validity as a measure of beginning teacher competency.

Table 3. TWS Sections linked to standards and accreditation requirements

	Performance-Based Standards for Colorado Teachers (PBSCT)	TEAC (& NCATE) Accreditation Standards
Unit Topic and Rationale	1, 2, 3, 4, 5, 6, 7	Subject Matter knowledge
Standards and Goals	1, 2, 3, 4, 5, 6, 7	Pedagogy
Instructional Plan	1, 2, 3, 4, 5, 6, 7	Pedagogy
Assessment	1, 2, 3, 4, 5, 6, 7	Pedagogy
Student Teaching Reflection	8	Caring Emergent Professional

Construct Validity: After running the factor analysis, three factors emerged from the item scores, where factor one was comprised of Standards, Goals and Assessment; factor two was comprised of Instruction and Management; factor three was comprised of Unit Topic and Rationale and Student Teaching Reflection. These results suggest that the components of the TWS are meaningful constructs that represent significant domains of teaching skills.

The results of the correlation between TWS components can be found in Table 4. Medium to low correlations indicate that scores between sections for each candidate are inconsistent.

Table 4. Correlations between TWS sections

	Unit Topic and Rationale	Standards and Goals	Instructional Plan	Assessment	Student Teaching Reflection
Unit Topic and Rationale	-				
Standards and Goals	.042	-			
Instructional Plan	-.037	-.530**	-		
Assessment	-.060	.690**	-.378**	-	
Student Teaching Reflection	.028	.101**	.069	.098**	-

**Correlation is significant at the 0.01 level (2-tailed)

RELIABILITY

Agreement statistics are typically measured by comparing pairs of rating for each score. CSU does not have this process in place, so an ANOVA was run to compare scores for each TWS section between university supervisors. Significant differences were found in two areas; Unit Topic and Rationale (significant differences between scorer #1 and all others) and Assessment (significant differences between scorer #1 and all others). These results indicate that there may be some inconsistency in scoring amongst the university supervisors.

DISCUSSION AND PLANS

The TWS was a meaningful assessment for the program. It connected teaching and learning, enabling candidates to combine and blend and interdependent set of processes, products, results and reflection. The analysis of data indicates that our teacher candidates demonstrate several competencies through their Teacher Work Samples. Overall, students met the following accreditation requirements:

- (1) Understand the subject matter they intend on teaching.
- (2) Utilize pedagogy that integrates theory and practice in efforts to meet the needs of diverse learners in real world educational contexts.
- (3) Are caring, competent emergent professionals.

Beyond accreditation requirements, further analysis of validity and reliability gave faculty a deeper understanding of how the TWS can be used to inform the pedagogical, programmatic and evaluative practices.

Because the sections of the TWS are all linked to CO performance based standards and TEAC and NCATE accreditation standards, the content of the instrument is valid. Through the factor analysis, each section of the TWS was included within the three factors which suggests that the components of the TWS are meaningful constructs that

represent significant domains of teaching skills. CSU teacher licensure is confident in using this instrument as a measure of teacher candidate preparation.

The correlations between the TWS sections among candidates were medium to low. Although there would appear to be inconsistency between the sections, students scored in the advanced range. While students are successful, the inconsistency between sections could be further examined. Faculty collaborative investigations could answer whether the variability is due to examiner practices or student success. This could strengthen the pedagogy and assessment practices throughout the program.

Training for scoring the TWS is provided for the university supervisors every semester. When rubric point values were changed in 2009, significant difference in scores from former years and 2009 indicated that the change was implemented with fidelity. This highlights the quality of the implementation of the modification.

Reliability is usually measured by comparing pairs of ratings for scores, and CSU does not have this practice in place. When comparing scores between university supervisors using ANOVA, one scorer was significantly different in the areas of Unit Topic and Rationale and Assessment. Although training is provided, we suggest a more formal process of baseline scoring and collaboration to improve assessment reliability.

The Teacher Work Sample is an assessment tool that also has much more data to be analyzed. Because the shift in teacher education has been moved to focus on outcomes (Levine, 2006), preparation programs need to examine their impact on student achievement. The National Network of Educational Renewal also strives to impact student achievement. The TWS at CSU assigns prospective teachers to examine student success in relation to their teaching. Section VI: Assessment Data and Analysis, requires student teachers to use student achievement data to analyze their instruction. We are implementing electronic TWSs in the coming semester. The storage of the TWSs will be organized so analysis of Section VI content will be feasible. This will give us a greater understanding of the impact our teacher candidates have on student achievement. Further analysis will also allow us to examine inter-rater reliability within the hierarchical structure of the university (i.e. tenure track faculty, clinical teachers, and university supervisors). Bullough's (2010) ethnomethodological study cautions the use of TWS outcomes for high stakes accountability and highlights that it can be problematic for faculty to come to consensus on TWS scores, which impacts inter-rater reliability. This will also be something for us to take into consideration in the future.

The discourse on teacher preparation calls for programs to defend what they do through proving their effectiveness. Program evaluation of this magnitude allows teacher preparation programs to determine strengths and weaknesses and inform their practice. Through increased accountability and effectiveness, candidates will be more fully prepared to enter the field of education. Focusing on teacher candidates' ability to demonstrate their effectiveness and adaptability toward meeting student needs is imperative. As the Cochran-Smith (2001) urges; "new teacher education ought to make room for discussions about outcomes that demonstrate how teachers know when and what their students have learned as well as how they manage dilemmas and wrestle with multiple perspectives" (p.180). Teacher Work Sample methodology proves to be a framework to meet this objective.

AUTHOR INFORMATION

Dr. Donna Cooner, School of Teacher Education and Principal Preparation, Colorado State University, Fort Collins, Colorado, USA 80523-1588. Dr. Donna Cooner is an associate professor at Colorado State University and Director of the School of Teacher Education and Principal Preparation. Her research interests include work in professional development schools, new teacher and principal induction, and program evaluation. She is also an author of *Becoming a Teacher in a Field-Based Setting*. E-mail: donna.cooner_gines@colostate.edu

Dr. Cerissa Stevenson is an assistant professor in the School of Teacher Education and Principal Preparation at Colorado State University. She is the Early Childhood Education Coordinator and her research interests include teacher and principal preparation. School of Teacher Education and Principal Preparation, Colorado State University, Fort Collins, Colorado, USA 80523-1588. E-mail: Cerissa.stevenson@colostate.edu

Dr. Heidi Frederiksen is a key advisor and coordinator of the principal licensure program in the School of Teacher Education and Principal Preparation at Colorado State University. Her research interests include teacher preparation and dispositions. School of Teacher Education and Principal Preparation, Colorado State University, Fort Collins, Colorado, USA 80523-1588. E-mail: Heidi.frederiksen@colostate.edu

REFERENCES

1. Bullough, JR., R.V. (2010). Proceed with caution: Interactive rules and teacher work sample scoring strategies, and ethnomethodological study, *Teachers College Record* 112 (3), 775-810.
2. Cochran-Smith, M. (2001). Higher standards for prospective teachers: What's missing from the discourse? *Journal of Teacher Education* 52(3), 179-181.
3. Cochran-Smith, M. (2006). *Policy, practices and politics in teacher education*. Thousand Oaks, CA: Corwin Press.
4. Colorado Department of Education (2000). Performance-based standards for Colorado teachers. Retrieved June, 2009 from http://www.cde.state.co.us/cdeprof/download/pdf/li_perfbasedstandards.pdf
5. Colorado Department of Education race to the top proposal (2010). Race to the top application for initial funding phase 2. Retrieved July 12, 2010, from Colorado Department of Education Web site: <http://www.colorado.gov/cs/Satellite/OIT-2/OIT2/1240228834570>
6. Darling-Hammond, L. (2005).
7. Darling-Hammond, L. (2006). Assessing teacher education: The usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education* 57(2), 120-138.
8. Darling-Hammond, L. (2006). Constructing 21st century teacher education. *Journal of Teacher Education* 57(XX), 1-15.
9. Denner, P. R.; Salzman, S. A.; & Harris, L. B. (February, 2002). Teacher work sample assessment: An accountability method that moves beyond teacher testing to the impact of teacher performance on student learning. Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, New York, NY.
10. Denner, P. R., Norman, A. D., Salzman, S. A., Pankratz, R., & Evans, S. E. (2004). The renaissance partnership teacher work sample: Evidence supporting score generalizability, validity, and quality of student learning assessment. In E. M. Guyton & J. R. Dangel (Eds.) *Research linking teacher preparation and student performance* (pp.23-55). Dubuque, IA: Kindall/Hunt Publishing Company.
11. Dewey, J. (1938). *Thought and language*. Newly revised and edited by Alex Kozulin. The MIT Press: Cambridge.
12. Duncan, A. (2009, Oct.). *Teacher preparation: Reforming the uncertain profession, remarks of Secretary Arne Duncan* at Teachers College, Columbia University. Retrieved July 9, 2010 from <http://www2.ed.gov/news/speeches/2009/10/10222009.html?exp=5>
13. Educational Testing Service. (2002). *A national priority: Americans speak on teacher quality*. Princeton, NJ: Author.
14. Feistritzer, E. (2004). *Alternative teacher certification: A state by state analysis 2004*. Washington, DC: National Center for Education Information.
15. Goodlad, J.I. (1990). *Teachers for our nation's schools*. San Francisco: Jossey-Bass Publishers.
16. Goodlad, J. I., Soder, R. & Sirotnik, K. A. (1990). *The moral dimensions of teaching*. San Francisco: Jossey-Bass Publishers.
17. Grossman, P.L., & Stodolsky, S.S. (1995). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher*, 24(8), 5-11, 23.
18. Hess, F. (2001). *Tear down this wall: The case for a radical overhaul of teacher certification*. Washington, DC: Progressive Policy Institute.
19. Jensen, J. W. & Winitzky, N. (1999, Feb.) What works in teacher education. Paper Presented at the Annual meeting of the American Association of Colleges for Teacher Education.
20. Ketter & Stoffel (2008). Getting real: Exploring perceived disconnect between education Theory and practice in teacher education. *Studying Teacher Education* 4(2), 129-142.
21. Levine, A (2006). Education school teachers: The education schools project. Retrieved July 15, 2010 from http://www.edschools.org/pdf/Educating_Teachers_Report.pdf.

22. Lorson, K. (2010). Making teacher work samples “work” for accreditation. *Journal of Physical Education, Recreation and Dance* 81(2), 54-57.
23. McConney, A. A., Schalock, M. D., & Schalock, H.D. (1998). Focusing improvement and quality assurance: Work samples as authentic performance measures of prospective teacher effectiveness. *Journal of Personnel Evaluation in Education* 11(3), 343-363.
24. Murray, F.B. (1996). *The teacher educator's handbook: Building a knowledge base for the preparation of teachers*. San Francisco: Jossey-Bass Publishers.
25. *The National Council for Accreditation of Teacher Education*. (2008). Retrieved October 28, 2009, from NCATE: <http://www.ncate.org/public/glossary.asp?ch=4#P>
26. National Foundation for the Improvement of Education. (1996). *Teachers take charge of their learning: Transforming professional development for student learning*. Washington, DC: Author.
27. National Research Council of the National Academies (2010). *Preparing teachers: Building evidence for sound policy. Executive Summary*. Washington, D.C.: The National Academies Press
28. Paige, R. (2002). Meeting the highly qualified teachers challenge: The secretary's annual report on teacher quality. Washington, DC: U.S. Department of Education, Office of Postsecondary Education.
29. Perry, C., Smith, D., & Woods McConney, A. (1998). Assessment framework for a proficiency-based teacher education program. Monmouth, OR: School of Education, Western Oregon University.
30. Schalock, M. D. (2002). Connecting teaching and learning: An introduction to Teacher Work Sample Methodology. In G. R. Girod, (Ed.). *Connecting teaching and learning: A handbook for teacher educators on teacher work sample methodology*. Washington, DC: American Association of Colleges for Teacher Education, (pp. 5-32).
31. TEAC. (2007). Retrieved November 28, 2009, from Teacher Education Accreditation Council http://www.teac.org/?page_id=170.
32. Vygotsky, L.V. (1986). *Thought and language*. Newly revised and edited by Alex Kozulin. MIT Press: Cambridge.
33. Wineburg, M. S. (2006). Evidence in teacher preparation: Establishing a framework for accountability. *Journal of Teacher Education*, 57(1), 51-64.
34. Wise, A. E., & Leibbrand, J. A. (2001). Standards in the new millennium: Where we are, where we're headed. *Journal of Teacher Education*, 52(3), 244-255.
35. Wood, S. N. (2000). Work sample methodology: Preparing teachers to assess their effectiveness in terms of student learning.
36. Wood, S. N. (2002). Summary, interpretation, and reflection in teacher work samples. In G. R. Girod, (Ed.). *Connecting teaching and learning: A handbook for teacher educators on teacher work sample methodology*. (263-286). Washington, DC: American Association of Colleges for Teacher Education.
37. Youngs, P., & Bird, T. (2010). Using embedded assessments to promote pedagogical reasoning among secondary teaching candidates. *Teaching and Teacher Education*, 26, 185-198.