

Fast and Reliable Evaluation of Preservice Teacher Electronic Portfolios

James Sulzen (james.Sulzen@uconn.edu)

Michael F. Young (myoung@uconn.edu)

University of Connecticut

Storrs, CT

Paper presented at the annual meeting of the
American Educational Research Association
in Chicago, IL, April 9-13, 2007.

Abstract

This study describes a rubric supporting fast and reliable assessment of preservice teacher electronic portfolios. The assessment calls for raters to quickly scan a portfolio to gain an overall impression, then dichotomously score a large number of indicators (e.g., educational philosophy, educational technology use, imaginative use of technology), followed by giving a score for the entire portfolio. Raters typically evaluated portfolios in 15-to-20 minutes and inter-rater reliability was 0.85, comparing quite favorably in speed and reliability in rating other complex student work, such as essays and term papers. Scoring a large number of dichotomous items for each portfolio provided a rater with a single coherent visual summary of a portfolio, which seemed to contribute to the reliability of the overall portfolio rating. Aggregating related indicators into subscale scores provided analytic measures of portfolio quality such as portfolio organization and technology skills. By utilizing indicators appropriate to a given portfolio's content and purpose, the technique described here is easily adapted to scoring portfolios from differing preservice teacher programs or scoring portfolios from different stages of a preservice teacher's educational career.

Table of Contents

Introduction.....	1
Developing the PSI240	3
Methods.....	6
Design	6
Participants.....	6
Student Teachers.....	6
Faculty Raters	7
Data Sources	8
Network Folder Based System Organization.....	8
TaskStream Commercial ePortfolio System.....	9
Portfolio Sampling.....	10
Procedures.....	11
Results.....	12
Discussion.....	16
Conclusion	19
References.....	21
APPENDICES	24
Appendix A - Portfolio Scoring Inventory (PSI240) Rubric	24
Appendix B – Score Record Sheet for Portfolio Scoring Inventory.....	27
Appendix C - Content Definition of Portfolios Used in the Present Study	28

Fast and Reliable Evaluation of Preservice Teacher Electronic Portfolios

Many teacher education programs are requiring students to create electronic portfolios (ePortfolios) of their work as a way of demonstrating aspects of technology, pedagogic, and professional competency (Anderson & DeMuelle, 1998; Batson, 2002; Delandshere & Arens, 2003; Lynch & Purnawarman, 2004; Strudler & Wetzel, 2005). Assessing these portfolios has largely been an unsatisfactorily addressed problem due to issues of validity, reliability, and the extensive time investment usually required to meaningfully assess the portfolio's qualitative content (Dollase, 1998; Herman & Winters, 1994; Wolf, Lichtenstein, & Stevenson, 1997; Wolfe & Miller, 1996). Wolf & Dietz (1998) characterized preservice teacher portfolios as primarily having a learning, assessment, or employment function. Teacher preparation programs with a portfolio requirement frequently use portfolios as a student capstone project, with capstone projects falling into the assessment category (even if they serve a learning or job search function) as it makes little sense to academically require a final project that has no meaningful consequences on the student's academic career. Yet the time pressures on students getting ready to graduate and the problems inherent in meaningfully assessing the portfolios leaves many preservice programs in a difficult position in terms of adequately fulfilling the assessment function (Delandshere & Arens, 2003; Strudler & Wetzel, 2005). This can leave the capstone project as more a 'hoop jumping' exercise than a valid learning or assessment experience as students rush to assemble their exit portfolios that nobody is going to seriously evaluate (Delandshere & Arens, 2003). Additionally, this largely administrative use of e-portfolios does not support the accountability function often expected of them (National Research Council,

2001; Interstate New Teacher Assessment and Support Consortium, 1992; National Council for Accreditation of Teacher Education, 2002).

While a number of researchers have relatively recently reported some success with reliable and valid scoring of preservice teacher portfolios (Burns & Haight, 2005; Denner, Norman, Salzman, & Pankratz, 2003; Yao, Foster, & Aldrich, 2006; Sulzen, Alfano, Zack, & Young, 2007), none reported on the usability of the portfolio scoring system or reported on the time investment required for evaluating each portfolio. Additionally, each of these reports has difficulties limiting the potential utility of the reported scoring system. Some of these systems require multiple raters to achieve acceptable reliability, further increasing the costs and lowering the system's pragmatic utility (Denner, Norman, Salzman, & Pankratz, 2003; Yao, Foster, & Aldrich, 2006) or have been tested only with portfolio raters who were also instructors of the assessed students and/or developers of the rubric (Burns & Haight, 2005; Sulzen, Alfano, Zack, & Young, 2007). A portfolio scoring system, to be practical, must produce reliable scores with raters unfamiliar with the students a rater is assessing, have validity, and be usable by people other than the scoring system's developers. While it is generally agreed that validity and reliability are achievable by careful selection of portfolio tasks, careful training of portfolio raters, and narrowly construing the judgments asked of raters (Moss, Sutherland, Haniford, et al., 2004; Herman & Winters, 1994), the scoring of portfolios is generally considered one of the most onerous and problematic aspects of portfolio implementation (Wolfe & Miller, 1996), making scoring perhaps the greatest challenge in implementing portfolios.

To address the above issues and have an effective way to evaluate developmental preservice teacher portfolios as a group, we created and assessed the validity and reliability of an electronic portfolio-scoring instrument that we call the PSI240. We designed the instrument to

allow for relatively fast scoring and to be easy to use, with a reasonable degree of reliability and validity. With the PSI240, a rater quickly scans a portfolio to gain an overall impression and then dichotomously scores a moderately large number of indicators that are signs of quality student performance. Based upon the dichotomous scores and the general impression, a rater assigns an overall score to the portfolio. The current study reports on the reliability and ease of use of this instrument. We reported companion data elsewhere concerning content, substantial, and structural validity (Messick, 1995) of the instrument (Sulzen and Young, 2004).

Developing the PSI240

To develop the instrument, we evaluated the portfolio assessment literature, particularly concerning teacher portfolio rubrics, and recommendations on teacher portfolio design and assessment (Barrett, 2001; Connecticut State Department of Education, 2004; Darling-Hammond, Wise, & Klein, 1998; Dollase, 1998; Durham & Bodzin, 2001; Goldsby & Faizal, 2001; Green & O'Sullivan Smyser, 1996; INTASC, 1992; ISTE-NETS, 2002; Martin-Kniep, 1998; NCATE, 2002; Walker, 2000). The portfolios developed by our students were really more “proto-portfolios” than real ones, developed in the first semester, under widely varying requirements across a number of instructors, and with only a few hours invested in them. As such, we adjudged it too complex to be able to produce an instrument with sufficient reliability and validity that could serve both grading and research purposes. These issues led us to design an instrument suitable for research on curricular interventions that we were then pursuing. We wanted an instrument that we felt would:

- Be targeted for use with raters who were experienced teacher educators;
- Be relatively easy and reliable for multiple portfolio raters to use;

FAST AND RELIABLE PORTFOLIO EVALUATION

- Highlight indicators of good portfolio construction such as we might find at an early stage of preservice teacher development;
- With the availability of exemplars and printed materials, require minimal rater training; and
- Be capable of easily adapting to changes in future portfolio assessment needs.

To address the above requirements, we chose to dichotomously score the presence / absence of some 17 criteria which we felt an experienced teacher educator could reliably detect and score (see Table 3 for the list of items we settled upon).

In the scoring scheme, raters were required to decide whether a portfolio contained any meaningful evidence of a particular criterion or not, such as presence of good navigation, some sort of philosophy of education, or discussion of some aspect of educational technology. As described below, these were developmentally early portfolios so we set the standard relatively low for scoring whether or not a portfolio contained evidence for a particular criterion. For example, for “educational background” a student would not receive credit for merely listing the names of high school and college attended, but would receive credit if the student additionally gave dates of attendance and identified the college degree program and expected date of graduation (see Appendix A for details of the rubric). Since items were dichotomously scored, no further credit was given regardless of how much more educational background may have been listed. Raters scored the other items similarly. The portfolio author did not need to place the particular datum under traditional headings - placing work experiences inside an essay on why the student wanted to become a teacher was just as acceptable as if listed under a more traditional heading. What mattered was whether there was evidence for a particular criterion no matter where the evidence occurred in the portfolio. However, we did recognize that some

students would likely go far and above their fellow students in at least some regards, but we could not predict the nature or areas we would likely encounter such exemplary work, nor know how best to take note of it in the scoring. As such, we added two other items (also dichotomously scored) to signify that a student had done something notable or had clearly spent far more effort than the assignment required. The seventeen criteria we settled on fell naturally into three subscales: professionally related work, technology items, and individuation and mechanical items (see Table 3). By looking for a relatively large number of criterion, but at a relatively low threshold of acceptability, we hoped to have a means of meaningfully measuring and differentiating what we expected to be a set of developing portfolios exhibiting a wide breadth, but limited depth of content.

In addition to the dichotomous scoring, raters provided a single overall subjective rating of the portfolio, scored 0 to 100, with an expected average / median of 75 corresponding to the somewhat traditional grading scheme. A score of “75” designated a portfolio for which the rater felt the student had done an adequate, though not particularly good job for the assignment (i.e., was “mediocre” corresponding in some sense to an ‘honest’ grade of “C”). In generating the overall score, raters considered the student work in terms of the portfolio resulting from a course assignment that was to become the basis for the student’s future ePortfolio, with the student using a pre-built template and easy-to-use on-line web page building tools.

After developing a draft of the PSI240, three content experts examined and provided feedback on the instrument, and we adjusted its content based upon their feedback. The authors then individually scored four randomly selected portfolios, consulted on scoring discrepancies, and made revisions to the scoring rubric and instrument to account for differences (see Appendix

A for the final rubric and Appendix B for the scoring sheet). Three of the initially scored portfolios were used as benchmarks to guide raters in subsequent scoring.

In retrospect, at the first scoring session among the authors, it might have been more effective to have jointly scored and revised the instrument rather than separately scoring and comparing our results post hoc. The post hoc discussion was probably not as illuminating as a joint scoring discussion most likely would have been, though it is unclear how a joint discussion would have changed the outcome.

Methods

This section describes the study design, participants, data sources, and procedures we followed to assess the reliability and utility of the PSI240.

Design

We conducted a generalizability study with two facets (rater by portfolio) by having four raters independently score eight randomly selected portfolios (Shavelson & Webb, 1991). This design provides a reasonable basis for screening the reliability of a rating instrument. For example, power for this design is 0.8 if the generalizability coefficient (rank-order reliability) is just 0.6; with a reliability of only 0.4, power is still 0.7 (Montgomery, 2001, p. 529). After the generalizability study, two of the raters scored an additional set of portfolios to assess subscale reliabilities.

Participants

Student Teachers

About 90% of the students entering the preservice program across two academic years agreed to participate in the study (see *Table 1*). These students were evenly divided between two

academic-year cohorts and were juniors in their first semester of our three-year integrated Bachelors / Masters program. The student make-up was over 95% Caucasian, predominantly from non-urban areas of Connecticut, with approximately two-dozen males in each year. All students learned to use the portfolio system and completed initial requirements for the portfolio in a required technology in education course in which they all enrolled. Each student enrolled in a section that met once a week with a dozen students per section. One of a half-dozen instructors taught each section, three of who also participated in this study as described below.

Table 1

Number of Student Participants and Scored Portfolios.

Cohort	Total Students in Cohort	Total Participants in Study (% of total)	Number of scored ePortfolios ^a (% of total)
1	131	121 (92%)	25 (19%)
2	123	112 (91%)	50 (41%)

^aRandomly selected for scoring.

Faculty Raters

Four raters participated in this study, each rater an experienced post-secondary instructor who was well versed in assessing undergraduate artifacts such as those contained in the portfolios. Two of the raters were familiar with the portfolio-scoring instrument (the authors) and the two other raters, both female, were not. Three of the raters (the two authors and one of the other raters) were instructors in the technology course in which the students developed the portfolios used in this study. However, as described in the Sampling section below, portfolio sampling precluded any instructor evaluating a portfolio from one of his or her own students. Using a mix of raters, blind to instructor and portfolio author identity, allowed us to generalize the results across a variety of raters who did or did not have prior familiarity with the rubric and evaluated performance based only upon portfolio content.

FAST AND RELIABLE PORTFOLIO EVALUATION

Data Sources

This study utilized portfolios created by the participants in the first semester of their teacher education program and as such, these portfolios were relatively early developmental efforts. While each cohort used a rather different portfolio development system as described below, the portfolio requirements and content were similar between the two years (see Appendix C). For these preliminary portfolios, we wanted to scaffold students in their initial portfolio organization and allow them to show their developing skills as a preservice teacher. Students used a template that identified the types of artifacts to provide. Some artifacts were information for the student to fill out in a template web page (educational and academic background) and others were assignments from specific courses, such as a lesson plans from their education courses or an essay describing their educational philosophy. Students could also upload artifacts of their own choosing.

While overall content was similar between the two years, each student cohort used a rather different on-line portfolio platform from each other. We briefly describe each portfolio system below to characterize the differences students experienced each year in developing their portfolios. The change in platforms afforded a more robust test of the portfolio assessment system described in this study than likely would have occurred had such changes not occurred.

Network Folder Based System Organization

With the first cohort, the school of education's information technology department provided a web server account for each student on a school file server. This, in essence, gave each student their own website with its own URL. These accounts allowed students to create folders with varying protection levels (public, self-only, or self-and-the-instructor). By storing HTML (i.e., web pages) and other browser-accessible files into their network folder, students

created web pages that were public, private, or could be restricted to instructor access. The use of student-managed websites had the advantage of using the existing technology infrastructure, but the disadvantage, but the disadvantage was that the students required significant instruction and technical support before the majority could use their accounts. Figure 1 shows a typical example of the first page of a student's portfolio.

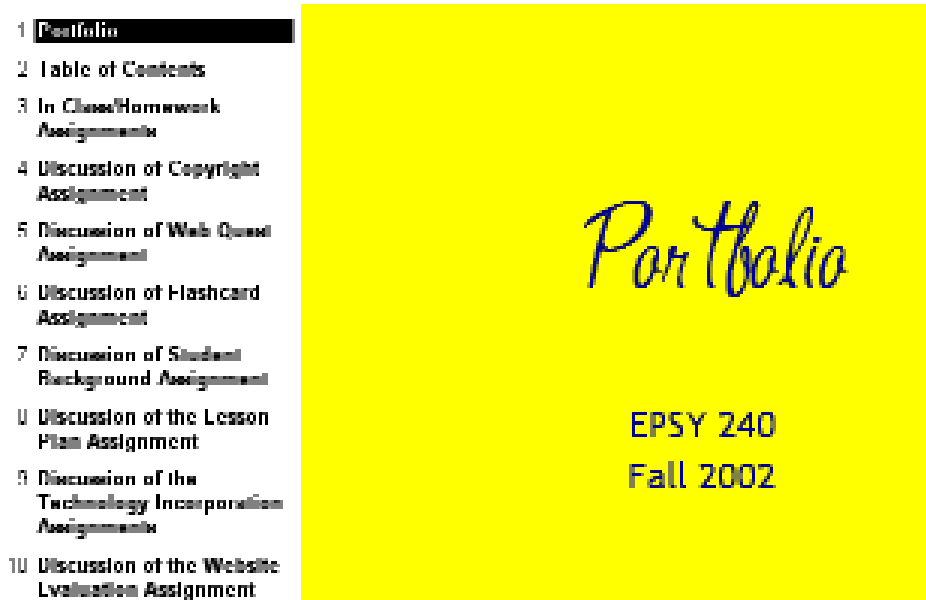


Figure 1

First page of a typical student portfolio produced with the web-folder ePortfolio platform.

TaskStream Commercial ePortfolio System

Due to many inherent limitations with the web-folder system just described, in the second year of this study and covering the second cohort, the school of education chose Taskstream as a relatively encompassing assessment system (<http://www.taskstream.com>). Taskstream is a web-based system providing a wide variety of on-line student, instructor, course, curricular, and teaching standards tools targeted particularly, though not exclusively, for a school of education environment. Students subscribed to the system on a yearly basis for \$40, approximately the cost

of a course text book, and used fill-in, pre-built project templates designed by instructors to create or upload portfolio artifacts such as their educational background, lesson plans, essays, and so on. For their actual portfolios, students used a standard template provided by the system (see Figure 2).

Authored by:
 Last modified: 10/30/2003 1:10:09 PM EST

Created with **TaskStream**
 www.taskstream.com THE TOOLS OF ENGAGEMENT

[Home](#) • [Educational Background](#) • [Teaching Experiences](#) • [Favorite Publications](#) • [Favorite Resources](#) • [Awards](#)

Bio



I am interested in teaching High School English. I am very interested in American and British Literature. I am currently enrolled as a student in the University of Connecticut's NEAG School of Education. I am currently an observer at Smith Middle School in Glastonbury in an English as a Second Language (ESOL) classroom as well as two different Language Arts classes. I am currently looking into the certification I would need to teach English as a Second Language in the future. When I graduate from the University of Connecticut I'd also like to coach Diving on the High School level in Connecticut.

File Attachments:
 1. [Picture](#)

Figure 2. First page of a typical student portfolio produced with the TaskStream Educator template.

Portfolio Sampling

Of the portfolios available (see *Table 1*), portfolios were randomly sampled in such a way that no rater evaluated portfolios authored their former or current students. Raters were also blind to the identity of the portfolio author's technology course instructor to preclude the potential biasing effects from such knowledge. However, perhaps 10% of the sampled portfolios did have content identifying the technology instructor, so this aspect of the sampling effort was not completely successful though we do not think it had significant impact on our results. Ensuring

raters were blind to acquaintance with the students or knowledge of the students' instructor limited the sample size, but allowed us to have greater confidence in the generality of our results. Specific samples for the generalizability and subscale reliability evaluations are described below.

Generalizability Study Sample: To maximize variance due to the portfolio facet, we used a stratified random sample of eight portfolios from the 2003-2004 cohort. The senior author informally reviewed 30 portfolios from course sections not taught by any of the faculty raters and classified each as low, or high quality. Four portfolios were randomly selected from each of the two groups to assure a wide range of portfolio quality in the generalizability study sample.

Subscale Evaluation Sample: To assess subscale reliabilities, after the generalizability study, an additional 75 randomly selected portfolios were scored with the rubric by one of two raters (the senior author and one of the non-author raters from the generalizability study). To check inter-rater agreement, 15 of the 75 portfolios were randomly selected and blindly scored by both raters, neither rater being aware of which portfolios were selected for double scoring; inter-rater correlation on the 15 portfolio item 18 portfolio score was $r=0.91$, $p<.01$. The total sample of scored portfolios (83 portfolios) allowed for an average of approximately five portfolios per item, considered a minimal requirement for a regression analysis of an instrument such as evaluated in this study (Russell, 2002). Since this study was a preliminary investigation, we felt this to be a reasonable trade-off between the effort involved and the information we obtained.

Procedures

Each of the four raters participating in the generalizability study scored the eight selected portfolios in a separate random order. The two female raters, neither of whom were previously familiar with the PSI240, used the instrument in a blind, untrained condition, working only from

the written materials in Appendices A and B, the three benchmark portfolios as scored examples, a simple verbal explanation of how to use the instrument, with no feedback on their judgments. For the subscale evaluation study, the two raters scored portfolios in a separate randomized order. Each rater was queried after participation as to total time spent rating portfolios and which was used to estimate the time effort required per portfolio.

Results

The results indicated that the scores from the overall PSI240 portfolio score (item 18) had a reliability of 0.85 (i.e., 85% of total variance due to differences among portfolios) and which are as good or better than results from similar studies and acceptable for research and general use (Burns & Haight, 2005; Denner, Norman, Salzman, & Pankratz, 2003; Netemeyer, Bearden, & Sharma, 2003; Yao, Foster, & Aldrich, 2006). Two of the three subscales also seemed potentially effective indicators of portfolio quality (the professional and technology scales), while the third subscale (mechanical and individuation) did not.

Table 2

Variance Components based on Item 18 Portfolio Score (Eight Portfolios by Four Raters).

Component	SS	df	Mean Square	F	Sig.	Variance Component	Reliability (fraction of total variance)
Rater	195.8	3	65.3	2.6	.08	5.0	2%
Portfolio	4803.2	7	686.2	27.1	.00	165.2	85%
Rater * Portfolio	531.4	21	25.3			25.3	13%

Table 2 displays the variance components calculated from a single-replicant two-way analysis of variance (ANOVA) with rater and portfolio as random factors and the summative portfolio score (item 18) as the dependent variable (Shavelson & Webb, 1991). As mentioned

above, reliability of the item 18 portfolio score as $G=0.85$ (95% confidence interval ranges from 0.72 to 0.91). The variance in portfolio score due to differences in portfolios themselves was highly significant, $F=27.1$, $p<.001$, $MSE=686.2$. This indicates that 85% of the variability in portfolio scores was due to differences in the portfolios themselves while the remaining 15% was due to inconsistencies between raters or in how raters viewed individual portfolios. The effect of rater approached significance, $F(3, 21)=2.6$, $p=.08$, partial *eta squared* =0.27, $MSE=25.3$, and which is some cause for concern that one or another rater may have consistently scored the portfolios higher or lower than the other raters. However, while there might be a statistically measurable effect of raters on the score, raters themselves only influenced score variability by 2%, which seemed acceptably small.

Figure 3 displays the item 18 portfolio scores from each of the four raters and which makes it apparent how consistent the raters were with each other. Each line in Figure 3 represents the rater scores for one portfolio. A perfectly flat line would represent perfect rater agreement and crossing lines indicate where raters disagreed on the relative quality of two portfolios. The high 0.85 item 18 portfolio score reliability reflects the relative flatness of and relatively few crossing lines apparent in Figure 3.

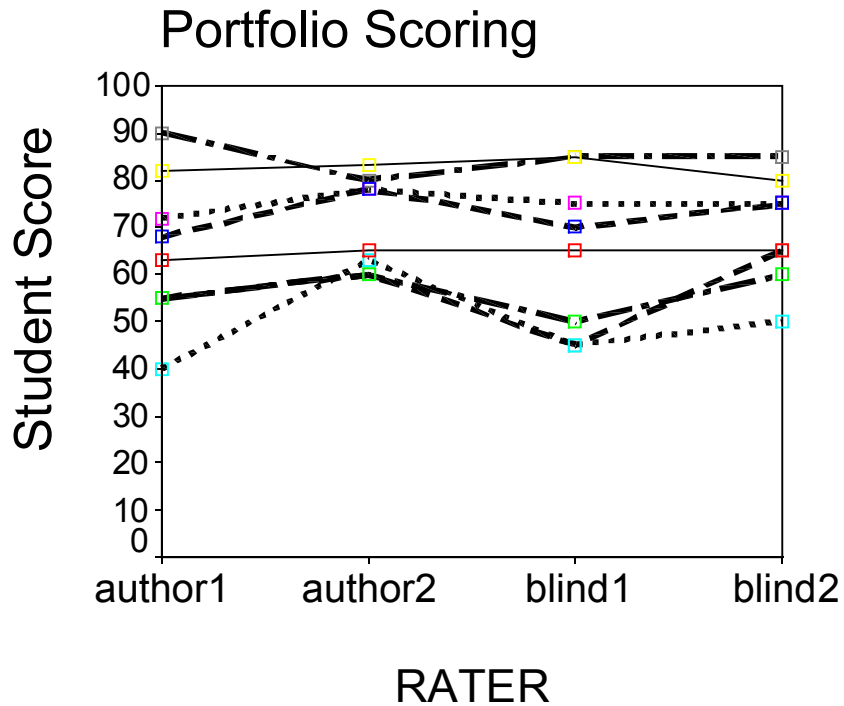


Figure 3. Rater Item 18 Portfolio Scores from Generalizability Study.

Note. Each line represents rater scores for one portfolio. Blind raters were blind to rubric development and only had printed rubric and scored benchmarks from which to work.

Table 3 lists the item statistics and scale reliabilities of the PSI240. For the analysis, items 14 and 15 were reverse scored to align them with the positive sense of the other items. Cronbach's alpha across all dichotomous items was 0.79 (see Table 3), which is considered quite good for research purposes, but is also expected with a large number of items (Netemeyer, Bearden, & Sharma, 2003). Three sub-scales that in theory might comprise more narrowly unidimensional subsets of the data were calculated and are listed in Table 3, labeled "Professional," "Technology", and "Mechanical and Individual." From examining the item statistics and initial reliabilities for each sub-scale, it was apparent that the raters inconsistently scored some items (items 3, 4, 9, 13-17). Consequently, the scales were reformulated as shown in the last column of Table 3 to create scales whose reliabilities exceeded 0.70 (Netemeyer, Bearden, & Sharma, 2003). This led to discarding the Mechanical and Individual scale which

had no significant theoretical justification, it being little more than a grouping of miscellaneous items.

Table 3

PSI240 Items and Scale Reliabilities (N=83 portfolios).

Item	Mean	SD	Reliability (based on items = x)	SMC ^a	Improved reliability if deleted ^b	Cronbach's alpha reliability (based on items marked x)
Overall Scale (items 1-17)	6.3	3.5	.80			.79
Professional Scale (items 1-8)	4.6	3.3	.54			.70
1 Clear opening	.81	.40	x	.48		x
2 Good navigation	.66	.48	x	.49		x
3 Educational & teaching goals	.22	.42	x	.14	.62	
4 Philosophy of education	.84	.52	x	.41	.56	
5 Educational background	.78	.42	x	.65		x
6 Professional bio info	.59	.50	x	.19		x
7 Reflective or self-evaluation	.59	.61	x	.44		x
8 Evidence of P-12 learning	.09	.30	x	.31		
Technology Scale (items 9-12)	1.1	1.6	.63			.79
9 Technology skills used in construction of portfolio	.50	.51	x	.21	.75	
10 Presence of educ. technology	.28	.52	x	.47		x
11 Wise integration of technology	.13	.42	x	.31		x
12 Tech. in service of pedagogy	.19	.40	x	.54		x
Mechanical & Individual Scale (items 13-17)	.47	.31	.43			<i>scale discarded</i>
13 Individuation / personalization	.53	.51	x	.18		
14 Spelling, grammar, & compositional flaws (scored 0 or -1)	-.03	.18	x	.04		
15 Noticeable technology flaws (scored 0 or -1)	-.03	.18	x	.04		
16 Bonus / something extra	.00	.00	x	-		
17 Clear extra effort / breadth / depth	.13	.32	x	.13		
18 Overall portfolio score (range 0-100, average/median ~75)	67.7	13.4				.88 ^c

^a SMC = Squared multiple correlation (proportion of variance of item in common with other items).

^b Blank cells in this column indicate item deletion would lower the subscale reliability.

^c Adjusted R squared when regressed on Professional and Technology scales.

In both the generalizability study and subsequent scoring, each portfolio required only 15 to 20 minutes on average to review and score with raters typically scoring three-to-four portfolios per hour.

FAST AND RELIABLE PORTFOLIO EVALUATION

Discussion

The PSI240 addresses a need for quickly assessing a large body of portfolios compiled by students each academic term in a preservice program. Average time to score each portfolio was 15 to 20 minutes, depending on the complexity of the student's work. Raters achieved this speed because the rubric encouraged them to follow a similar systematic scanning process with each other and with each portfolio that did not require a rater to review every page and every line of student work. Raters merely needed to identify the types of artifacts that existed and that each met the rubric's standards. The scoring speed compared favorably with scoring other involved student work such as multi-page essays or term papers. Additionally, the scoring speed occurred with good inter-rater agreement, had good face validity, and could potentially provide meaningful evaluation of student work. If portfolios are to fulfill their promise in teacher education, there need to be ways to evaluate portfolios in a fast but reasonably reliable manner, even if only on a relatively surface review basis, or else they risk becoming a pedagogical device to which educators pay lip service, but are just too onerous to regularly and meaningfully evaluate.

Historically, portfolio evaluation programs have achieved high reliability with extensive rater training lasting upwards of a week or more (e.g., Connecticut State Department of Education, 2004; National Board for Professional Teaching Standards, 2002) or apprenticed inexperienced raters with more experienced ones (Yao, Foster, & Aldrich, 2006). However, given the limited resources in teacher education institutions, competing demands on those likely to be responsible for the portfolio evaluation function (teacher education faculty), and the need for year-to-year consistency, preservice teacher portfolio evaluation needs to avoid extensive training or retraining to be practical. It is important that a realistic system is usable more or less

‘as is’ by most teacher educators. The generalizability study reported in this article employed four teacher educators, two of who developed the evaluation instrument, but two “outside” raters who used the instrument with no training, had to learn to use it through printed materials and scored on-line exemplars. The simplicity of the judgments the rubric called for was probably fundamental to the consistency of the raters. However, the consistency of performance of the outside raters with that of the rubric authors is notable, and we think important hallmark, of our portfolio evaluation approach.

We have no specific data, but it is our supposition that the organization of the scoring sheet’s list of scored indicators, provided a rater with a visually coherent summary of a portfolio that seemed to contribute to the reliability of the overall portfolio rating. We organized the scoring sheet to specifically provide a visually compact overview of each portfolio side-by-side with the rater’s previously scored portfolios (see Appendix B). Organizing a rater’s accumulated judgment data in such a fashion seems an important element contributing to a rater’s improved self-consistency. The relatively quick scoring process also contributed by allowing raters to more readily remember and mentally compare the quality and scores across portfolios than a slower scoring system would have afforded.

The Professional and Technology scale scores had good reliability and seem potentially useful for obtaining research results or providing feedback on the effectiveness of curricular interventions. These scales indicate that the cursory examination process used by raters seems capable of providing potentially meaningful information in addition to an overall evaluation of the portfolio.

The scoring technique used in this study, rapid impressionistic dichotomous evaluation of many indicator items, is probably unsuitable for grading student work for no other reason than its

unconventionality. Perhaps a more significant difficulty for grading purposes is the need to limit the number of rubric items and the relatively low ‘ankle high’ criteria required to positively score items. Good grading practices call for fully informing students of grading criteria. Informing students of the specific rubric indicators would likely motivate them to focus solely on these, ignoring deeper and broader, but still important matters. However, the issue of how to inform students about scoring criteria is comparable to similar issues in paper and pencil test item construction. Test item selection presumes the individual test items are ‘sampling’ student knowledge and skills. We generally do not inform our students of which individual items to expect on a test, but instead give general criteria or areas in which we expect students to be capable. Portfolio content definition and communication of criteria should be similar in this way to more traditional forms of student testing. However, students are informed of expected portfolio content definition and grading criteria, we feel there is still an open question as to the suitability of our scoring system for actual grading of student work due to the facile surface-level judgments called for in scoring, even allowing for the system’s seeming reliability and validity.

Selecting rubric items for a portfolio evaluation rubric as advocated in this study, while similar in some ways to paper and pencil test item selection, does add an additional error component due to rater judgment that is not typically present in standardized tests, but of course still exists in open-ended or essay answer items. With portfolio assessment, this potential error component is inescapable since we are interested in evaluation of complex performances that require inherent rater judgment. The issue is not so much of whether or not to involve raters, but, instead, to select items that insures construct validity and insures acceptable minimization of variance due to the rater and rater/performance interaction (Messick, 1996). It is an open question whether the assessment technique described here affords more sophisticated portfolio and

preservice teacher performance than that of the early developmental portfolios used in this study. More substantive indicators than used in this study's rubric should make for a more substantial assessment. For example, in an assessment we are currently working upon, we have found the following to be reliably scored indicators in a student teacher's portfolio regarding general pedagogic knowledge (Sulzen, in preparation): use of social and constructivist practice, repertoire of teaching models, activity-based and interactive, and differentiation of instruction. Similarly, we have found raters reliably consistent in scoring the following for evaluating instructional delivery skills: appropriate timing and pacing, clarity of directions, active learning during presentations, and facilitating whole class in dialog. Such indicators are much more substantive than the ones used in this study, and extending the portfolio assessment technique described here seems feasible for more sophisticated preservice teacher work.

Conclusion

The reliable and efficient evaluation of preservice teacher electronic portfolios is a challenging problem. This article discusses an approach that while perhaps unconventional, provides a means for a researcher or a teacher education institution to affordably gather data about preservice teacher portfolio content and quality. Raters were very consistent with each other in using much of the rubric described in this study and added limited variance to overall portfolio score (see Table 2). It was possible to form reliable sub-scales from groups of rubric items that we had expected to be related, allowing the rubric to be somewhat analytic as well as summative in form. Portfolio raters required very little training support; they quickly scanned and quickly scored each portfolio, spending 15 to 20 minutes on each, making the process relatively affordable in terms of time use.

FAST AND RELIABLE PORTFOLIO EVALUATION

The portfolio assessment technique described here requires development of a rubric consisting of relatively simple portfolio content and criteria items, both of which that meaningfully sample preservice teacher performance. Each item needs to be sufficiently straightforward and intuitive for a suitably experienced teacher educator to easily score. The simplicity of items means raters do not require extensive training or need to invest much time in scanning and scoring a portfolio since each individual item judgment is easily made. As in quick-answer test construction, the effectiveness of this technique depends upon the evaluation rubric having a sufficient number of consequential items to meaningfully sample the range of performance expected and to allow for some degree of inconsistency among raters. This means, as always, the devil, per force, is in the details of item selection in terms of creating a meaningful instrument.

This portfolio evaluation instrument and instrument design addresses to some degree the need to quickly and effectively assess a large number of developmental ePortfolios, a task most observers consider should be done at least once a year and preferably more often, but which is considered one of the most onerous aspects of portfolio use (Delandshere & Arens, 2003; Wolfe & Miller, 1996). While perhaps not suitable for grading preservice teacher performance in a thoroughgoing summative manner, we have used PSI240 in making data-driven decisions regarding an academic program and in deciding upon the relative effectiveness of differing pedagogical approaches (Sulzen & Young, 2004).

References

- Anderson, R. S. & DeMeulle, L. (1998). Portfolio use in twenty-four teacher education programs. *Teacher Education Quarterly*, 25(1), 23-31.
- Barrett, H. (2001). Electronic Portfolios. To be published in *Educational Technology - An Encyclopedia*. Retrieved July 30, 2003, from <http://helenbarrett.com/portfolios.html>
- Batson, T. (2002). The Electronic Portfolio Boom: What's It All About? *Syllabus*, 16(5)
- Burns, M. K. & Haight, S. L. (2005). Psychometric properties and instructional utility of assessing special education teacher candidate knowledge with portfolios. *Teacher Education and Special Education*, 28(3/4), 185-194.
- Campbell, D. M., Melenyzer, B. J., Nettles, D. H., & Wyman Jr., R. M. (2000). *Portfolio and performance assessment in teacher education*. Boston: Allyn & Bacon.
- Connecticut State Department of Education (2004). BEST teacher portfolio performance rubrics. Retrieved Feb. 20th, 2004, from <http://www.state.ct.us/sde/dtl/t-a/best/portfolio/rubrics.htm>
- Darling-Hammond, L., Wise, A. E., & Klein, S. P. (1998). *A license to teach: Building a profession for 21st Century schools*. San Francisco: Jossey-Bass.
- Delandshere, G. & Arens, S. (2003). Examining the quality of the evidence in preservice teacher portfolios. *Journal of Teacher Education*, 54(1), 57-73.
- Denner, P. R., Norman, A. D., Salzman, S. A., & Pankratz, R. S. (2003). *Connecting teaching performance to student achievement: A generalizability and validity study of the Renaissance Teacher Work Sample assessment*. Paper presented at the Annual Meeting of the Association of Teacher Educators (Jacksonville, FL, February 17, 2003). (ERIC Document Reproduction Service No. ED474079).
- Dollase, R. H. (1998). When a state mandates portfolios: The Vermont experience. In Lyons, N. (ed.), *With portfolio in hand*. NY: Columbia University.
- Durham, C. S. & Bodzin, A. (2001). Scoring for preservice teachers' electronic portfolios: Issues of feasibility and reliability. In SITE 2001 PT3 Papers: *Proceedings of Society for Information Technology and Teacher Education International Conference* (12th, Orlando, Florida, March 5-10, 2001).
- Goldsby, D. & Faizal, M. (2001). Now that your students have created web-based digital portfolios, how do you evaluate them? *Journal of Technology and Teacher Education*, 9(4), 607-16.
- Green, J. E. & O'Sullivan Smyser, S. (1996). *The teacher portfolio*. Basel, Switzerland: Technomic Pub.

- Herman, J. L., & Winters, L. (1994). Portfolio research: A slim collection. *Educational Leadership*, 52(2), 48-55.
- International Society for Technology in Education (2002). International Society for Technology in Education - National Education Technology Standards for Teachers (ISTE-NETS*T). Retrieved April 1, 2004, from http://cnets.iste.org/teachers/t_stands.html
- Interstate New Teacher Assessment and Support Consortium (1992). Model standards for beginning teacher licensing, assessment, and development. Retrieved July 22, 2004 from http://www.ccsso.org/Projects/Interstate_New_Teacher_Assessment_and_Support_Consortium/
- Lynch, L. L., & Purnawarman, P. (2004). Electronic portfolio assessments in educational and instructional technology programs: Are they supporting teacher education?. *TechTrends*, 48(1), 50-58.
- Martin-Kniep, G. O. (1998). Why am I doing this? Purposeful teaching through portfolio assessment. Portsmouth, NH: Heinemann.
- Messick, S. (1996). Validity in performance assessments. In G. W. Phillips (Ed.), *Technical issues in large-scale performance assessment* (NCES 96-802) (pp. 1-18). Washington DC: National Center for Educational Statistics.
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist*, 50(9), 741-749.
- Montgomery, E. C. (2001). Design and analysis of experiments, 5th Edition. NY: John Wiley & Sons.
- Moss, P. A., Sutherland, L. M., Haniford, L., Miller, R., Johnson, D., Geist, P. K., Koziol, S. M., Star, J. R., & Pecheone, R. L. (2004). Interrogating the generalizability of portfolio assessments of beginning teachers: A qualitative study. *Educational Policy Analysis Archives*, 12(32), 1-70.
- National Board for Professional Teaching Standards (NBPTS) (2002). What teachers should know and be able to do. Retrieved July 22, 2004 from <http://www.nbpts.org/standards/stds.cfm>
- National Council for Accreditation of Teacher Education (2002). Professional standards for the accreditation of schools, colleges, and departments of education. Retrieved July 28, 2004 from http://www.ncate.org/2000/unit_stnds_2002.pdf
- National Research Council (2001). Testing teacher candidates: The role of licensure tests in improving teacher quality. Washington, D.C.: National Academy.
- Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). Scaling procedures. Newbury Park, CA: Sage.

- Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin*, 28(12), 1629-1646.
- Shavelson, R. J., & Web, N. W. (1991). Generalizability theory: A primer. Newbury Park, CA: Sage.
- Strudler, N., & Wetzel, K. (2005). The diffusion of electronic portfolios in teacher education: Issues of initiation and implementation. *Journal of Research on Technology in Education*, 37(4), 411-444.
- Sulzen, J., & Young, M. F. (2004). Fostering web and multimedia skills of preservice teachers via electronic portfolios. Paper delivered at the annual meeting of the American Educational Research Association, San Diego, CA, April 14-18, 2004.
- Sulzen, J., Alfano, M., Zack, J., & Young, M. (2007). Validation of eFolios for preservice teacher education. Paper presented at the American Educational Research Association annual conference Chicago, IL April 9-13, 2007
- Walker, S. (2000). Electronic portfolios in assessment of preservice teachers. WebNet 2000, San Antonio, TX October 30 - November 4, 2000.
- Wolf, K., & Dietz, M. (1998). Teaching portfolios: purposes and possibilities. *Teacher Education Quarterly*, 25(1), 9-22.
- Wolf, K., Lichtenstein, G., & Stevenson, C. (1997). Portfolios in teacher evaluation. In J. H. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practice*. Thousand Oaks, CA: Corwin Press.
- Wolfe, E. W. & Miller, T. R. (1996). Barriers to the implementation of portfolio assessment in secondary education. *Applied Measurement in Education*, 10(3), 235-251.
- Yao, Y., Foster, K., & Aldrich, J. (2006). Generalizability study of a team-based scoring approach. Paper presented at the American Educational Research Association annual meeting in San Francisco, CA, April 7-11, 2006.

FAST AND RELIABLE PORTFOLIO EVALUATION

APPENDICES

Appendix A - Portfolio Scoring Inventory (PSI240) Rubric

This appendix lists the PSI240 rubric. Item numbers in the Criteria section correspond to elements on the scoring sheet listed in Appendix B.

General Notes

- A. The main purpose of this instrument is to be able to meaningfully differentiate the 2003 first-semester preservice portfolio / educator-web pages from each other and similarly to be able to differentiate the comparable 2002 ePortfolio project (the web folder based ones from last year).
- B. The minimal EPSY240 assignment required students use the TaskStream “Educator’s Biographical” template to begin creating an ePortfolio. This template included a separate web page for each of the following: Home / opening page, educational background, courses/classes taught, favorite publications, favorite resources, and awards. As such, the ‘standard you should expect’ is about what you would expect of students who are building their initial portfolio as part of a 1-unit EdTech course.
- C. Each of the major headings below should be read as “Evidence of...”. Evidence should extend past a mere pro forma statement, but should be a relatively low threshold given that these are first-semester, one-unit artifacts.
- D. Scoring will probably be mostly 0/1 for each numbered item below (with provision of a “2” for a truly outstanding exemplar). A score of “2” should be annotated with rationale. A “2” is something you would point to and tell everyone else that THIS is absolutely one way how it should be done – something you would not expect to see except on a professional product.
- E. Subheadings below are mostly clarification details.
- F. Some of the categories below might appear somewhat redundant. The intention is to use the scores in some yet undetermined weighting scheme to establish ultimate assessment measures.
- G. Certain score items below are likely (or not) to score a “1” (or “0”) for every portfolio on TaskStream because TS “provides” the feature for free (i.e., good opening and navigation); the items exist to support using this scoring system for non TaskStream-based portfolios so as to assess divergent scoring validity of this instrument.

Criteria

- 1) Clear opening
 - 1.1 Title, introduction/orientation, and perhaps presence of a TOC
 - 1.2 Is it apparent from the opening that one is looking at some sort of portfolio (preferably that of an educator or want-to-be educator) and how one would likely find relevant information, assuming one had a specific interest.
- 2) Good navigation
 - 2.1 Something other than a linear “page turner”; user friendly
 - 2.2 Logical and reasonably accurate grouping of linkages; effective TOC
 - 2.3 Multiple blank or missing screens or links that clearly do not appropriately connect with what is expected
- 3) Educational and teaching goals
 - 3.1 Must be something more than “I want to teach history in High School” or whatever; if stated in such generic terms than score zero.
 - 3.2 Goals must be personalized and specific to individual. Example: I want to teach kids how history directly influences our every day experiences and so they see the direct relevance of history to them.
- 4) Philosophy of Education
 - 4.1 Educational/teaching/learning philosophy/theory is stated or very apparent
 - 4.2 The philosophy / theory does not need to be “correct”(i.e., textbook) and certainly not compelling;
 - 4.2.1 Should at least be reasonable (for a first-semester I/BM student);
 - 4.2.2 Must be beyond just pro forma, not be vapid nor just plain wrong.
 - 4.2.3 Must have some degree (and need not be much) of meaningful content.
- 5) Educational background
 - 5.1 List of schools & programs attended; possibly, courses taken;
 - 5.2 Should be something more than just pro forma
 - 5.3 Should provide an individualized background of student’s education
 - 5.4 Statement of high school and colleges attended with major subjects (or other supporting detail) is sufficient.
- 6) Professional biographical information

FAST AND RELIABLE PORTFOLIO EVALUATION

- 6.1 Description of relevant work & teaching experience, management of kids, meaningful description of relevant educational preparation, etc.
- 6.2 Significant and specific detail about current or past professional and/or educator experience and capabilities or about current preparation leading to expected future professional capabilities.
- 7) Reflective or self evaluation
 - 7.1 Captioning: What each element is, why it is present, and what it is evidence of.
 - 7.2 Reflective, self-evaluative or insightful essays or commentary
- 8) P-12 student learning
 - 8.1 Evidence that actual K12 student learning took place because of the portfolio owner's individual efforts
 - 8.2 Examples: Classroom photos, lesson plans taught with, descriptions of teaching experiences, assessments performed, example student work, etc.
- 9) Technology skills used in construction of portfolio
 - 9.1 Imaginative or unusual/unique exploitation of the technical capabilities that extends beyond typical naïve use
 - 9.2 Exploitation of the technology or tools beyond basic word processing, copy/paste, or fill-in-the-blank web form skills.
 - 9.3 Examples include the use of: Non-trivial HTML; hierarchically organized web pages; screen captures; Java/JavaScript, or other scripting capabilities; custom digitized media; creation of animated GIFs.
- 10) Presence of educational technology
 - 10.1 Mention of any technology in connection with an educational context
 - 10.2 Must identify the technology and provide justification or function of the technology in an educational context.
 - 10.3 This is strictly for "hard" technology items and does not include anything that falls under PedTech
 - 10.4 Example: "use computers to take notes"
- 11) Wise Integration of Technology
 - 11.1 The use of technologies proper in a wise and intelligent manner that improves the education in a way that a comparable and simpler non-technology methodology would not. Lesson plan, lesson plan concepts, or examples making useful/meaningful use of technology integration.
 - 11.2 As with everything else, this item should have a relatively low threshold, but the use of the technology must not be a pro forma or gratuitous reference, but be intelligently relevant in context (e.g., "using GIS(Geographic Information Databases) to teach time" does not cut it).
 - 11.3 This is strictly for "hard" technology items (such as EdTech / WebTech) and does not include anything that falls under PedTech
 - 11.4 Example: Foreign language learners using ePals, email, iVisit, etc. to interact with native language speakers ; simulation programs to support lesson content; word processing to revise multiple drafts or reformat for multiple educational purposes. Use of word processor to draft and polish letters to CEOs about the importance of rain forests to our ecology.
 - 11.5 Non-examples: Word processing to take notes; non-educative / irrelevant uses of email; using the web to look up info readily available in non-technology sources; gratuitous technology use.
- 12) Technology in service of pedagogy (PedTech)
 - 12.1 Evidence of knowledge of PedTech (problem-based, collaborative, wide variety of technology available to tackle educational problems, ethical & social issues of technology in education, etc.). This item addresses aspects of technology that are educationally relevant but do not fall under the auspices of the prior items.
 - 12.2 This covers things which are not necessarily pedagogical or technological in nature, but which tend to come up when one brings technology into an educational environment.
 - 12.3 Examples: Mention that one should avoid gratuitous tech use; tech particularly suited to constructivist, collaborative, problem-oriented approach; social & ethical issues of tech; tech should be used in service of educational problem solving; copyright; fair access / digital divide; assistive usage; suggestions for classroom management of limited technology resources (computer allocation, grouping students to use computer together, etc.).
- 13) Individuation
 - 13.1 Personalization differentiating the portfolio from others and individually reflecting the author
 - 13.2 Significant esthetic design elements
 - 13.3 Design, visual, or other elements that are out of the ordinary or serve to differentiate from other portfolios
 - 13.4 Examples: Educationally or reflectively appropriate poems or personal stories; well-crafted and customized overall portfolio;
- 14) Any spelling, grammar, and other compositional difficulties (score 0 or -1)
 - 14.1 I usually let one minor spelling flaw go
 - 14.2 However, the great bulk of the portfolios seem virtually flawless, so I'm inclined to be rather tight on this standard
- 15) Any noticeable technical flaws (score 0 or -1)
 - 15.1 Broken links, missing graphics, obscured text, or other obvious deficiencies that are technically correctable;
 - 15.2 You should be careful in assessing link rot that is not the fault of the portfolio author
 - 15.3 Ignore what seem to be browser-specific errors
 - 15.4 I usually let them have at least one minor technical flaw (such as a broken/rotten link) before marking them for this
- 16) Bonus / Something extra
 - 16.1 Examples: significant and relevant set of web resources, books, particularly telling vignettes, outstanding technical capabilities demonstrated, etc.

FAST AND RELIABLE PORTFOLIO EVALUATION

- 16.2 There's a good chance that scoring a "2" in any other category will generate a "1" in this category. A "2" in this category would require a set of outstanding elements in the portfolio
- 17) Clear extra effort / breadth / depth
- 17.1 Some meaningful effort and thought apparent in the preparation of the portfolio itself (as opposed to attaching lots of work from other courses) and the effort clearly customizes the content specifically to the individual
- 17.2 Advanced: far exceeds an expectation for assignment
- 18) Overall score (0-100, average/median is ~75)
- 18.1 Used for internal and inter-rater reliability assessment
- 18.2 Base score upon what is reasonably expected of a first-semester "portfolio" of a preservice teacher for a one-unit (or perhaps a pair of one-unit courses).
- 18.3 Average/median score for 2003 TaskStream portfolios should be about 75; the bulk of scores (1 SD?) should probably fall in the 60-85 range.
- 18.3.1 Examples:
 - 65: All elements are listed, but not necessarily present; the portfolio / web pages are basically a disappointment and would never be used for their content or for any purpose (other than homework fulfillment requirements); student invested little mental and creative effort in fulfilling assignment; little or no content beyond pro forma; there are one or two missing (or basically missing) required elements.
 - 75: Student filled out every required element (beyond just pro forma on over half the items), with some personalization; did not include any meaningful work from other courses or from outside the bounds of the required assignment. Overall, a mediocre product.
 - 85: Student included significant and meaningful information for each required element; included at least one (preferably at least two) significant elements not required, but very relevant to a portfolio; portfolio is definitely well personalized and is perhaps a basis for starting a real portfolio.
- 19) Item counts
- 19.1 A) Distinct web pages or screens
- 19.2 B) Meaningful web links or web sites related to education or to the bio
- 19.3 C) Graphic items
- 19.4 D) File attachments
- 19.5 E) Word count of narrative text constructed specifically for the portfolio (estimated?)
- 20) CT DOE BEST Rubrics
 - Note: Following are taken from the major categories of the CT BEST Elementary Ed rubric and are tentative and probably not germane to the great majority of the ePortfolios. These are included to distinguish student work that seems to significantly exceed the "first-semester, one-unit course" context and extended into containing true professional portfolio-like evidence. A "2" in one of these categories says the evidence is at least comparable to what you would expect a professionally produced teacher ePortfolio to provide.
- 20.1 A) Content / subject-specific knowledge
- 20.2 B) Pedagogical knowledge / skills
- 20.3 C) Instructional design
- 20.4 D) Instructional implementation
- 20.5 E) Assessment knowledge / skills
- 20.6 F) Analyzing teaching & learning
- 21) Comments / Rationale
- 21.1 Short explanation of why any particular "2" was scored; anything else of note
- 21.2 Put a letter in the box and write a similarly marked note below

FAST AND RELIABLE PORTFOLIO EVALUATION

Appendix B – Score Record Sheet for Portfolio Scoring Inventory

Each column in the form below records scores for one portfolio. Score an item according to the rubric listed in Appendix A.

Scorer: 3/14/2004 Date portfolio last modified: _____ Student name: _____

MEANINGFUL EVIDENCE OF:					
1	Clear opening				
2	Good navigation				
3	P Educational & teaching goals				
4	R Philosophy of education				
5	O Educational background				
6	F Professional biographic info				
7	Reflective or self evaluation				
8	P12 student learning				
9	T Technology skills in construction				
10	E Educational technology				
11	C Wise integration of EdTech/WebTech				
12	H Pedagogy in service of technology				
13	O Individuation				
14	T Any spelling/grammar problems (0 / -1)				
15	H Any noticeable technical errors (0 / -1)				
16	E Bonus - something special about this				
17	R Clear extra effort				
18	Overall score (75 +/-)				
19	Item counts				
A.	C Screen Count				
B.	O Web Links / Web Sites				
C.	U Graphic Items				
D.	N File attachments				
E.	T Word Count				
20	CT DOE BEST Rubrics				
A.	Content / subject-specific knowledge				
B.	Pedagogical knowledge / skills				
C.	Instructional design				
D.	Instructional implementation				
E.	Assessment knowledge / skills				
F.	Analyzing teaching & learning				
21	Comments/Rationale				

(put a footnote # in the box and write your comment at bottom of page)

Appendix C - Content Definition of Portfolios Used in the Present Study

This appendix describes the first semester portfolio requirement for the participants. We deliberately made the assignment relatively unstructured and open-ended, intending to start students in a long-term portfolio construction endeavor and leaving room for each student to embellish or not as she or he saw fit. Not surprisingly, there was a wide variety of quality across the portfolios produced.

Portfolio Project

Most of the portfolio comes from your on-going course work (i.e., just copy/paste your weekly assignments to your portfolio as you go). The final portfolio preparation should be little more than preparing a table of contents and short introduction about the portfolio or about yourself, adding some comments to it about each item (the portfolio sketches), and making sure all the links work. If you like, you may certainly go beyond these guidelines and extend the project by say adding a navigation bar, an index, include work from other classes, include deeper reflection pieces than we have asked for, etc. Fundamentally, for this course, the portfolio is intended to just get you started and to require little more than a compilation of your course homework assignments.

Portfolio Project Grading Rubric

40% Content: Inclusion of each homework assignment (not including final lesson plan project): (40 – All homework included and complete; 30 – Most homework included or several items are markedly incomplete; £ 20 – Substantial homework items missing or are incomplete.)

20% Portfolio Sketches: Short description (preferably only one or two sentences) for each homework assignment or portfolio item describing its purpose and/or how it fits in the portfolio. These should be included in a logical way (such as in the table of contents or as part of an introductory section) so that a reader can quickly decide what they would want to look at.

40% Organization: Organization and navigation with properly functioning links and other organizational elements as appropriate to assist the reader in accessing the document. (40 – Title page and introduction, table of contents, fully functioning links, and clear organization. 30 – Missing or weak organizational elements, some links not functioning, etc. 20 – None or substantially missing organizational elements, many links not functioning, or other major mechanical problems.)