International Journal of ePortfolio
http://www.theijep.com

2012, Volume 2, Number 2, 187-197
ISSN 2157-622X

Improving the Quality of Evidence-Based Writing Entries in Electronic Portfolios

David W. Denton
Seattle Pacific University

The problem investigated in this study was whether entries written to an electronic portfolio by preservice teachers improved in quality after an intervention was deployed. The study also compared portfolio metadata to writing quality scores to determine whether there was a relationship. Participants included a convenience sample of 11 undergraduate students enrolled in a teacher education program. Primary analyses focused on comparing portfolio entries, written before and after the intervention, using a repeated measures design. Secondary analyses involved calculating correlations between writing quality and portfolio metadata. Results showed that writing improved at a statistically significant level, $t(10) = 4.99$, $p < .001$, $d = 3.16$, 95% CI = 1.91 to 5.00. In addition, statistically significant correlations were found between writing quality and the number of unique terms shown on portfolio tag clouds, $r = .60$, $N = 11$, $p < .05$, $d = 1.50$, as well as writing quality and the total number of portfolio entries, $r = .72$, $N = 11$, $p < .05$, $d = 2.08$. These findings suggest that the intervention improved writing quality on entries made to electronic portfolios and that metadata predicted the quality of portfolio content.

A standard represents something that is widely accepted by authorities, or consensus, and it is used to make comparisons or shape behavior. Since the 1990s, educational reformers have organized their efforts around standards as a way to promote educational equality (Meier, 2000; Urban & Wagoner, 2009). Teachers and administrators in K-12 settings have been contending with standardized assessments, published standards, and more recently, Common Core Standards since this time.

Teachers and students in higher education have not escaped the momentum of education reform (Hill-Jackson & Lewis, 2010; Moss, 2007). The initiatives first encountered in primary and secondary classrooms are manifesting themselves as reworked accountability systems for students in post-secondary settings. For example, according to Basken (2007), hundreds of colleges are using standardized student-achievement data to allow for comparisons between institutions. These colleges are also posting performance-related data on their web sites to promote transparency.

Whether in K-12 or higher education settings, a significant feature underlying many of these accountability systems is evidence-based learning (Millett, Payne, Dwyer, Stickler, & Alexiou, 2008), which consists of soliciting student thinking on a particular problem, building consensus around a solution, and gathering evidence to support the solution and compare its outcome to a standard (Eitel & Steiner, 1999). The focus on evidence, comparison, and standards constitutes the core set of ideas upon which accountability and evaluation systems are predicated. However, another facet, arguably more student-centered and aligned with principles of folio thinking, is using evidence-based learning as a way to promote achievement through repeated cycles of reflection and assessment. Eitel and Steiner (1999) identify this analytical process as the "plan-do-check-act-cycle" (1999, p. 510).

Evidence-based learning is a model which is particularly useful to programs that train students to adopt and use professional practices and standards, such as teachers, nurses, doctors, lawyers, engineers, and clergy (Carnegie Foundation, n.d.; Scharton, 2013). Teacher education is a fitting example since education students are subject to national, regional, and state accountability initiatives. Specifically, students enrolled in teacher training programs engage in evidence-based learning as a way to demonstrate competency on professional knowledge and skills and to complete requirements for licensing (Cambridge, Cambridge, & Yancey, 2009). Along with showing a positive impact on achievement, these models also fulfill accreditation requirements, which are applied to licensing programs through multiple oversight organizations (Ewell, 2006, Schechter, 2007). For example, the National Council for Accreditation of Teacher Education (NCATE) requires that teacher candidates demonstrate knowledge of content, pedagogy, and the profession, along with showing the necessary dispositions for helping all students learn (National Council for Accreditation of Teacher Education, 2012).

Moreover, teacher candidates are required to master the broader standards established for the particular college or university that they attend. For example, the Association of American Colleges and Universities (2012) has indicated that graduates be prepared for twenty-first century challenges by learning how to think critically and creatively, along with showing continuous improvement of their communication skills.

The purpose of these systems, whether deployed within K-12 or post-secondary settings, is to promote

accountability and improve learning outcomes. One way that policymakers and researchers organize these systems is around curriculum, instruction, and assessment (Pellegrino, 2006; Porter & Smithson, 2001). However, governing bodies, such as state departments of education, have typically invested more time and resources in developing standards and assessments, which tend to govern instructional methods (Resnick, Rothman, Slattery, & Vranek, 2004).

Nevertheless, accountability and competency requirements for, and measures of, teachers in training have caused schools of education to integrate an evidence-based learning approach with additional procedures for recording and showing preservice teacher knowledge and skills (Barrett & Knezek, 2003; Wilhem et al., 2006). For more than 20 years, paper-based portfolios were the leading method for displaying this type of evidence (Ayan & Seferoglu, 2011; Barton & Collins, 1993). Shulman (1998) defined portfolio as a "structured documentary history of a (carefully selected) set of coached or mentored accomplishments substantiated by samples of student work and fully realized only through reflective writing, deliberation, and serious conversation" (p. 23).

Evidence-Based Learning and Assessment

As part of this evolution, advances in technology and increasing accreditation requirements, such as the need to track student progress on performance of standards over time, have made electronic portfolio systems more appealing to program administrators (Pecheone & Stansbury, 1996; Smith, Cook, Faulkner, & Peers, 2007; Zeichner & Wray, 2001). One reason for this is that training programs are required to show that exiting students have mastered specific knowledge and skills for employment (Yao, Aldrich, Foster, & Pecina, 2009). As a result, many schools of education have adopted portfolio systems, sometimes referred to as digital portfolios, e-folios, ePortfolios, or web-folios, to take advantage of efficiencies and technological enhancements, such as searching, retrieving, changing, linking, and organizing narratives and artifacts (Strudler & Wetzel, 2005).

Defining Electronic Portfolio

According to Abrami and Barrett (2005), an electronic portfolio is "a digital container capable of storing visual and auditory content including text, images, video and sound" (p. 2). Electronic portfolios contain many of the same features as paper-based portfolios. However, there are several advantages to electronic portfolios, including the ability to create hyperlinks, show metadata, and share contents efficiently with multiple viewers (Abrami & Barrett, 2005).

Wilhelm et al. (2006) distinguishes between two types of electronic portfolio platforms, including generic tools (e.g., WordPress and Blogger) and customizable vendor tools (e.g., TaskStream and LiveText). Both generic and vendor-based portfolio systems are implemented as an approach to evidence-based learning and assessment design, where the principle goal is to align standards to products and processes (Millett et al., 2008). Learning focused on generation and analysis of evidence is based on the identification of desired outcomes. In teacher education programs, desired outcomes are typically aligned with lists of knowledge and skills, also known as professional standards. These standards are authored by professional organizations or more likely, they are dictated by government legislation (Millett et al., 2008). Once professional standards are established, then system designers identify the types, and amount of evidence necessary for showing mastery. Mastery is demonstrated as portfolio authors engage in evidence-based learning, specifically through the presentation of claims and arguments which are connected to professional standards and artifacts through logical explanation.

Assessing evidence-based learning depends upon the complexity of the claims, standards, and evidence, along with consequences associated with results, such as the decision to license an education student for service. Since licensing is a significant decision, sophisticated evaluation systems are necessary to ensure the validity and reliability of evidence-based assessment results (Millett et al., 2008). Although Yao et al. (2008) found that electronic portfolios were insufficient for showing validity and reliability of teacher competence, most researchers agree that they do provide efficient and stable repositories of evidence-based learning (Abrami & Barrett, 2005; Smith et al., 2007; Wilhem et al., 2006). In addition, electronic portfolio authors are able to show a wide range of evidence, as suggested by Abrami and Barrett (2005), while also enabling them to synthesize this evidence into complex arrays of performance data. Synthesized performance data is one of the factors that contribute to valid and reliable assessments (Mislevy, Almond, & Lukas, 2004).

Writing and Portfolio Content

Although electronic portfolios enable authors to include a broad array of data, the most common type of content is writing. This writing appears in different formats, such as descriptive, analytical, and reflective. As a result, assessing an electronic portfolio also means simultaneously assessing the quality of the writing

shown on it. Assessment systems that require written responses will necessarily involve the evaluation of writing skill (Millett et al., 2008). One result of this relationship is that the writing shown on electronic portfolios can be a source of construct-irrelevant variance, especially if the purpose of the portfolio is to assess professional knowledge and skills (Mislevy et al., 2004). Alternatively, if performance standards are aligned with writing, or perhaps more broadly with communication, then writing ability will decrease as a source of construct-irrelevant variance. Although professional teaching standards do include elements of communication, in teacher education the emphasis tends to focus on profession-specific abilities, such as parent and community involvement, planning for instruction, classroom management, and so on.

Whether communication skills are acknowledged or ignored as an element of electronic portfolio use, there is clearly a relationship between the quality of writing used to construct portfolio entries and portfolio performance. Carney (2006) and Yao et al. (2009) have suggested investigating this relationship, specifically in the area of teacher education. Nevertheless, searching the Educational Research Information Center database for the terms "electronic," "portfolio," and "evidence" produced five current studies published in peer-reviewed journals that applied experimental or quasi-experimental methodology. Each of these studies examined writing assignments, aligned with principles of evidence-based learning, presented in an electronic portfolio format.

Previous Research Relating to Electronic Portfolios and Evidence-Based Learning

Although the analysis of previous research is limited in its scope, primarily as a result of the small number of studies which met search criteria, it does suggest some trends. One of these is that writing assignments focused on evidence-based reasoning and presented in electronic portfolios have a positive effect on preservice teacher knowledge and skills (Ayan & Seferoglum 2011; McIntyre and Dangel, 2009; Shepherd & Hannafin, 2011). However, there is limited consensus about this. For example, two studies, by the same principal author, suggest that the correlation between portfolio performance and measures of teacher competency were weak or nonexistent (Yao et al., 2008; Yao et al., 2009). Another trend is that studies tended to use qualitative methodology, such as observation, interview, and self-report data. Likewise, most studies included small sample sizes ($N < 10$). Finally, most of the studies used vendor-based portfolio systems, such as Taskstream and LiveText, although one study reported using WordPress blogs as a portfolio container. While this final trend does not address factors associated with writing quality and portfolio performance, it does have implications for how writing is assembled and presented.

With regard to assessing knowledge and skills, McIntyre and Dangel (2009) examined the effects of having six preservice elementary teachers complete electronic portfolio assignments based on teaching standards during a semester of internship. According to results collected from observations and interviews, participants reported that portfolio assignments increased their knowledge of professional standards and provided a method for showing growth in teaching ability.

In a similar study, Yao et al. (2009) examined the effects of deploying electronic portfolio assignments on preservice teacher knowledge and skills by collecting and analyzing interview data. Participants included eight preservice teachers. Interview data showed that participants perceived that portfolio assignments improved their capacity for reflection. However, participants also reported that the portfolio was not an accurate representation of teaching competence.

Another study by Yao et al. (2008) involved analysis of nearly 200 preservice teacher portfolios using quantitative measures. Results showed that portfolio performance was not correlated with other metrics of teaching ability. For example, Yao et al. found that portfolio scores were not predictive of standardized tests scores associated with general aptitude or subject matter competency, such as the ACT, C-Base, or Praxis II. However, portfolio scores did show a statistically significant relationship when compared to grade point average (Yao et al., 2008).

Ayan and Seferoglu (2011) analyzed the contents of portfolio entries from eight undergraduate preservice teachers during a semester of student teaching. Similar to the findings of Yao et al. (2008), who suggested that portfolio assignments were not predictive of teaching competency, results showed that participants wrote descriptions twice as often in comparison to analytical or evaluative compositions. Nevertheless, results from interview data also showed that participants believed that regular portfolio assignments helped them analyze their beliefs about instruction and classroom-based decision making (Ayan & Seferoglu, 2011).

Contrasting the study by Ayan and Seferoglu (2011), Shepherd and Hannafin (2011) designed an intervention incorporating specific portfolio writing instructions, including question prompts, assignment directions, and technical support materials. The study by Shepherd and Hannafin (2011) involved six participants, who completed three extensive assignments, consisting of four parts each. For these three assignments, participants responded to multiple questions about lesson planning and analysis and examination of student evidence. Results from

interview data showed that explicit instructions on portfolio assignments improved participants' ability to examine evidence and write plans for improving their instruction (Shepherd & Hannafin, 2011). However, participants reported needing individualized support and coaching due to the sophistication of the assignments (Shepherd & Hannafin, 2011).

With the exception of the study by Yao et al. (2008), which calculated correlations, and the study by McIntyre and Dangel (2009), which reported mean portfolio scores, all of the results in this summary were derived through qualitative methodology. Moreover, none of the studies used pretest-posttest methods of comparison or analysis of metadata. One possible explanation for the absence of metadata is that researchers from three of the six studies were using vendor-based portfolio systems, which are unable to generate this kind of information. Ayan and Seferoglu (2011) did report using WordPress, which is a blog platform that shows metadata through tag clouds and archives; however, these factors were not analyzed.

Lastly, none of the studies, except for the one by Shepherd & Hannafin (2011), examined interventions related to writing. This is notable since writing is an important method for presenting electronic portfolio content (Abrami & Barrett, 2005; Shulman, 1998).

Research Questions

The summary of research suggests that additional studies are needed to examine the relationship between writing quality and portfolio performance using quantitative measures, such as pretest and posttest methods, along with analysis of metadata. The purpose of the study described here was to operationalize evidence-based writing, investigate the relationship between writing quality and portfolio performance, and determine the effects of a writing intervention on the quality of teacher candidates' electronic portfolio entries. Specifically, research questions for this study included the following:

1. What is the relationship between writing quality and portfolio performance?
2. How do writing interventions affect entries written to electronic portfolio?
3. What is the relationship between writing quality and electronic portfolio metadata?

Method

Context of Study

The participants for this study consisted of a convenience sample of undergraduate students enrolled in a teacher preparation program. Participants included 11 seniors, 10 females and one male, who had completed 10 weeks of a 20-week teaching internship. These participants were scheduled to graduate within three months from the time that the study took place. Six participants were earning endorsements in elementary education, two in physical education, two in elementary special education, and one in secondary special education. The group consisted of participants from European descent only. The mean grade point average for the sample was 3.43.

Participants had created electronic portfolios nine months before the beginning of the study, using WordPress blogs. Each portfolio showed a landing page, or blog page, along with four auxiliary pages showing professional teaching standards. Each portfolio showed a tag cloud and archive.

Participants began writing entries to their portfolios at an average rate of one every two weeks. The contents of portfolio entries varied. For example, some described instructional theory, presumably written for a specific course; while other entries recounted events based on classroom observations. These entries were assessed by course instructors and practicum supervisors using a variety of methods, such as comments and points.

Primary Measure

Three portfolio entries were scored for each student using a repeated measures design. Two of these entries were written by participants and assessed by the instructor before the intervention was deployed. The oldest entry, further referred to as the first entry, was written nine months before the intervention was administered. The next entry, further referred to as the second entry, had been written between one day and one month before the beginning of the intervention. The third entry was written during intervention. The intervention lasted one hour, spread across two class sessions. Class sessions were separated by one week. Participants wrote, and then revised, their third entry outside of class. The first, second, and third entries were scored for writing quality. Writing quality was operationalized using a rubric, further referred to as the writing quality rubric (see Appendix).

This rubric contained five columns and two rows. Columns were scaled from 0 (*deficient*) to 4 (*exemplary*). The first row assessed the integration of artifacts used to show evidence of teaching competence. Artifacts included lesson plans, student work samples, teaching videos, or course papers, among other items. To achieve a score of 2 or above on this criterion, participants had to reference the artifact and interpret or evaluate its impact on their practice or student learning. The second row assessed the participants' analysis and evaluation of their teaching in comparison to a given professional standard. To achieve a score of 2 or above

on this criterion, participants had to reference the professional standard, analyze and evaluate their performance in comparison to the standard, identify significant conclusions about their teaching practice, and support their conclusions by referencing the artifact.

Descriptors, or cells, for each level of performance on the rubric were taken from VALUE assessments (AACU, 2012). The rubric was evaluated for reliability by calculating a Pearson correlation coefficient based on data taken from five observations made by two scorers, for three participants, comparing the first and third entries. Results showed a correlation of .82.

Primary Tests

Two paired sample t-tests were conducted to analyze differences between the writing quality of the first and third entry and the second and third entry. Although this study involved a relatively small sample size, calculating a large Cohen's *d* effect size, .80 or above, at an alpha level of .05, with three measures, required a statistical power of at least .93 (Faul, Erdfelder, Buchner, & Lang, 2009). Results from paired sample t-tests showed effect sizes well above the .80 threshold.

In addition, descriptive statistics showed that scores on the second and third entry were sufficiently distributed for parametric analysis, with skewness and kurtosis values below 1 (see Table 1). However, this was not the case for the first entry. The uniform distribution of results on the first entry were the outcome of participants scoring either one point or no points on the writing quality rubric, $M = .55$.

Secondary Tests

Participants' third entry was also scored according to a four-point scale of reflective writing designed by Kember, McKay, Sinclair, and Wong (2008). The original Kember et al. (2008) scale identified four levels of reflective writing according to letter designations, including (a) *critical reflection,* (b) *reflection*, (c) *understanding*, and (d) *nonreflection*. These letters were assigned numerical values from 4 to 1, respectively, and then used to assess the third entry ($M = 3.18$, $SD = .41$).

In addition, two types of metadata were collected from student portfolios before implementing the intervention. Each portfolio showed a tag cloud and archive (Figure 1). The number of words or phrases, further referred to as terms, in each tag cloud were counted ($M = 18.36$, $SE = 10.42$). Course numbers and generic titles were excluded (e.g., "EDU 1234," "weekly blog," "entry #4"). The total number of portfolio entries was also counted by summing from the numerals shown on each portfolio's archive menu ($M = 23.00$, $SE = 9.92$).

Scores from another portfolio assignment, which were not used to answer the research questions for this study, showed positive correlations with the number of tag cloud terms and total portfolio entries. The correlations were statistically significant (mean $r = .64$, $N = 11$, $p < .05$), indicating some convergent validity between writing quality and metadata.

Three Pearson correlations were calculated between writing quality scores for the third entry and the following predictor variables: (1) level of reflective writing defined by Kember et al. (2008), (2) tag cloud term count, and (3) total number of portfolio entries.

Intervention

The writing intervention included the following instructional practices: (1) explicit direction on content and format, (2) communication of assessment criteria, (3) evaluating evidence, (4) instructor and peer feedback, and (5) revising. These practices were deployed as participants began writing their third portfolio entry. Graham and Perin (2007) identified these methods as characteristic of the following approaches to writing instruction: procedural facilitation, product goals, inquiry, feedback, and process writing. According to Graham and Perin (2007), these instructional practices have a positive impact on writing skill and writing quality.

The intervention began with participants reading a short list of instructions, which provided a general description of the electronic portfolio writing assignment. This activity lasted approximately five minutes.

Following this, the instructor showed the assignment rubric to participants and identified its six criteria and four-point scale. The assignment rubric was similar to the writing quality rubric in two ways. First, each rubric indicated that participants were to cite a specific program standard and to write content that showed competence on this standard. Citing the program standard meant that participants were to identify the complete teaching standard, either verbatim, or to define it in their own words. Participants were instructed to organize the content of their writing around this standard. Second, each rubric indicated that participants were to show and reference an artifact. Showing and referencing an artifact meant that participants were to support their written conclusions with evidence. Evidence could include images, attachments, and screenshots of student work samples, lesson plans, and videos or pictures of teaching, among other items. However, the assignment rubric showed four additional criteria including word count,

Table 1
Descriptive Statistics for First, Second, and Third Entries

Entry	*M*	*SD*	Skewness	Kurtosis
First	.55	.52	.21	2.44
Second	2.36	1.70	.82	.88
Third	5.82	1.08	.43	.83

Figure 1
Example Portfolio Tag Cloud and Archive

Note. Terms in the tag cloud increase in size depending on the number of times the term is associated with an entry. The numbers in parentheses on the Archives menu show the number of entries submitted that month.

mechanics, on-time submission, and citation of an authoritative source, such as an article or textbook.

After the instructor discussed writing expectations and presented the assignment rubric, participants had five days to compose their entries. At the conclusion of the five days, participants submitted their entries for formative feedback from the instructor. Participants accessed feedback online through a learning management system. Although the instructor used the assignment rubric to generate feedback, a numerical score was not assigned. Rather, the instructor indicated (1) how the entry met or exceeded the criteria of the assignment rubric, and (2) specific areas of weakness needing revision.

During the next class session, after submitting the third entry and receiving feedback online, participants read their entry aloud to a peer with a paper copy of the assignment rubric nearby. Participants then discussed the entry using the rubric as a guide. Listening peers made suggestions for improvement. Participants switched reading aloud and listening roles and repeated the process.

After this activity, which took approximately 15 minutes, the instructor asked participants to revise and resubmit their entries. One day later, the instructor scored revised entries using the writing quality rubric.

Results

Primary Test Results

A paired sample t-test was conducted to evaluate whether the quality of participants' third portfolio entry improved in comparison to the second entry. Results indicated that the mean score for the third entry (M = 5.82, SD = 1.08) was significantly greater than the

mean score for the second entry, $M = 2.36$, $SD = 1.70$, $t(10) = 4.99$, $p < .001$. The standardized effect size index, d, was 3.16, with some overlap in the distributions for rubric scores between the second and third entries, as shown in Figure 2. The 95% confidence interval for the mean difference between the two ratings was 1.91 to 5.00.

A second paired sample t-test was conducted to compare differences between the writing quality of the third and first entries. Results indicated that the mean score for the third entry was significantly greater than the mean score for the first entry, $M = 0.55$, $SD = .52$, $t(10) = 22.24$, $p < .001$. The standardized effect size index was 14.07, with no overlap in the distributions for rubric scores between the third and first entries, as shown in Figure 2. The 95% confidence interval for the mean difference between the two ratings was 4.75 to 5.80.

Secondary Test Results

A Pearson correlation was computed to assess the relationship between scores for the writing quality of the third entry and scores assessing the level of reflection in written work (Kember et al., 2008). Results showed a statistically significant correlation, $r = .77$, $N = 11$, $p < .01$, $d = 2.41$.

A second correlation was computed to assess the relationship between the third entry and tag cloud term counts. There was a statistically significant correlation, $r = .60$, $N = 11$, $p < .05$, $d = 1.50$. A final correlation between third entry writing quality scores and the total number of portfolio entries was calculated, and it also showed a statistically significant result, $r = .72$, $N = 11$, $p < .05$, $d = 2.08$.

A summary of these findings, displayed as a correlation matrix, is shown in Figure 3.

Discussion

Relationship Between Writing Quality, Portfolio Performance, and Writing Intervention

The quality of student writing improved significantly in comparison to entries written before the intervention. This finding corroborates research by Graham and Perin (2007) who found that writing interventions aligned with procedural facilitation, product goals, inquiry, feedback, and process writing, improved participants' writing skill and writing quality. However, writing quality in this study was defined according to characteristics of evidence-based learning, specifically integration of artifacts and evaluation of teaching in comparison to a given professional standard. In addition, participants' third entry showed reflective writing, characterized by descriptions of theory and practice, practicum experiences, and personal insights about teaching (Kember et al., 2008). Similarly, studies by Ayan and Seferoglum (2011), McIntyre and Dangel (2009), and Shepherd and Hannafin (2011) also found that entries written to electronic portfolios improved participants' awareness and understanding of professional standards.

Alternatively, Yao et al. (2008) found portfolio scores correlated with grade point average, but not other measures, such as standardized tests associated with general aptitude or subject matter knowledge. Likewise, results from this study showed correlations between the scale developed by Kember et al. (2008), which assessed levels of reflective writing, and the writing quality of the third entry. However, these are largely measures relating to writing skill, with content focused on professional standards, analysis of evidence, and reflective composition. Whether quality portfolio entries predict real professional effectiveness is an important question, but it is also a question outside the scope of this study. Nevertheless, since portfolio entries correlated with the Kember et al. (2008) scale, there is at least some indication that the level of reflection shown on portfolio entries changed as a result of the intervention.

Relationship between Writing Quality and Electronic Portfolio Metadata

The metadata analyzed in this study, including the number of unique terms in a tag cloud and archives showing the total number of entries, was predictive of the writing quality of the third entry. Notably, the total number of portfolio entries was a stronger predictor of writing quality in comparison to the number of terms shown on a tag cloud. These results suggest that metadata is useful to instructors as an informal assessment measure of the general writing quality of electronic portfolio entries. However, including metadata as an electronic portfolio feature only appears to be available through generic tools, such as WordPress and Blogger.

Limitations

This study included three notable limitations. The sample size was small, and represented mostly females. However, studies by Ayan and Seferoglu (2011), McIntyre and Dangel (2009), and Shepherd and Hannafin (2011) involved sample sizes with less than 10 participants and also included mostly females. In addition, the principal investigator was also the instructor, which can lead to reactivity bias (Slavin, 1992). Nevertheless, educational studies, involving investigators who deploy interventions, are not uncommon (Dignath & Büttner, 2008). For example,

Figure 2
Boxplots of Writing Quality Rubric Scores for the First, Second, and Third Entries

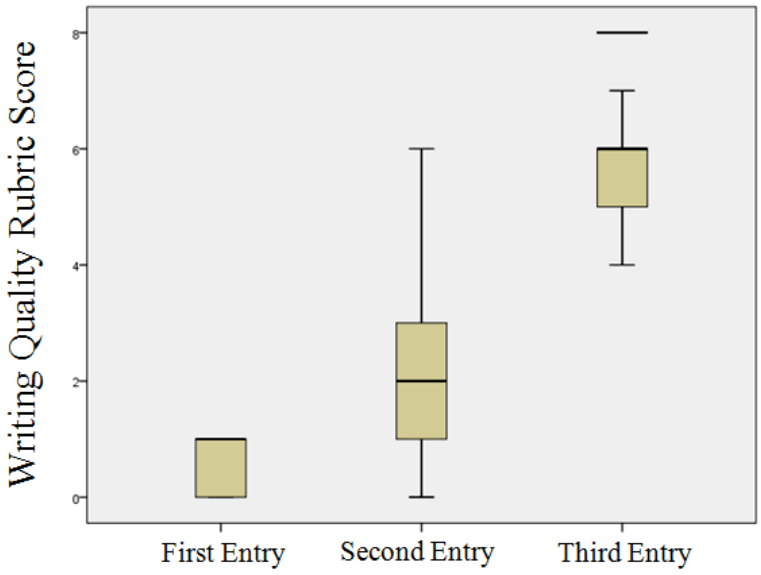
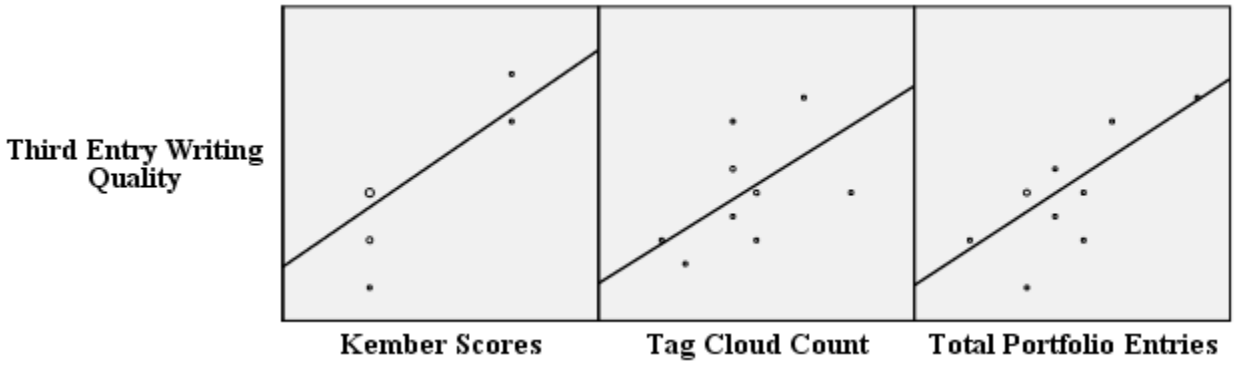


Figure 3
Correlation Matrix



Note. The matrix shows statistically significant relationships between scores for the writing quality of the third entry, level of reflective writing (Kember et al., 2008), tag cloud term count, and total number of portfolio entries.

Jenson (2011) acted as the instructor and investigator in a study examining participants' self-regulation and use of electronic portfolios, without identifying the potential for reactivity as a limitation. Finally, this study used a narrow definition of writing quality, which was operationalized using specific criteria, derived from characteristics of evidence-based learning and VALUE assessments (AACU, 2012). However, these characteristics were specifically chosen to assess written entries made to electronic portfolios, focused on the ability of preservice teachers to reference evidence, integrate teaching standards, and write meaningful conclusions about their practice.

Conclusion

One implication from this study is that electronic portfolio assignments designed to assist students with writing promotes outcomes aligned with specific evaluation criteria. In addition, rather than assigning a reflection with little or no direction, which students tend to define in different ways (Gustafson & Bennett, 2002), instructors may improve the results of written entries by (1) giving explicit direction on content and format, (2) communicating the assessment criteria, (3) requiring evaluation of evidence, (4) providing feedback, and (5) permitting revisions. However, these strategies may not be appropriate for soliciting open-

ended and exploratory responses in the context of problem-based learning or ill-structured settings. Rather, these strategies are more likely to serve instructors in their efforts to align portfolio content with specific professional standards or evaluative criteria.

A valid argument related to the use of electronic portfolios for purposes of alignment is that it reduces opportunities for authors to engage in reflection and self-actualization. This is a fair criticism. Indeed, Barrett and Knezek (2003) cautioned against using electronic portfolios exclusively for assessment purposes. Nevertheless, professional training, especially teacher education programs, are using electronic portfolios to satisfy certification and licensing requirements. As a result, some examination of how portfolios can be used to accomplish both student-centered and profession-centered goals is warranted.

Finally, since electronic portfolios are used for purposes that extend beyond student-centered learning and growth, it is likely that instructors will increasingly adopt practices to align portfolio content with system-defined outcomes, at least in professional education settings. An important question related to this is whether the instructional practices, useful for alignment, have a permanent or temporary effect on the quality of portfolio entries once support is withdrawn. More broadly, there is research to suggest that electronic portfolios are being partially subsumed into education reform efforts. These efforts emphasize a finite set of ideas, such as evidence, comparison, and standards, along with accountability and transparency. It is important to determine if integrating reform-based practices with electronic portfolios has as real and lasting effect. Research related to either of these questions is sure to be informative. Perhaps more importantly, investigating these questions will assist teachers and students in managing education reform efforts as they exert an influence on how electronic portfolios are used.

References

Abrami, P. C., & Barrett, H. C. (2005). Directions for research and development on electronic portfolio. *Canadian Journal of Learning and Technology, 31*, 1-15.

Association of American Colleges and Universities (AACU). (2012). *Valid assessment of learning in undergraduate education*. Retrieved from http://www.aacu.org/value/

Ayan, D., & Seferoglu, G. (2011). Using electronic portfolios to promote reflective thinking in language teacher education. *Educational Studies, 37*(5), 513-521. doi:10.1080/03055698.2010.539782

Barrett, H. C., & Knezek, D. (2003, April). *e-Portfolios: Issues in assessment and preservice teacher preparation*. Paper presented at the Annual Conference of the American Educational Research Association, Chicago, IL. Retrieved from http://electronicportfolios.com/portfolios/AERA2003.pdf

Barton, J., & Collins, A. (1993). Portfolios in teacher education. *Journal of Teacher Education*, *44*(3), 200-210. doi:10.1177/002248719304400307

Basken, P. (2007). A year later, spellings report still makes ripples. *Chronicle of Higher Education, 54*(5).

Cambridge, D., Cambridge, B. L., & Yancey, K. B. (2009). *Electronic portfolios 2.0: Emergent research on implementation and impact*. Sterling, VA: Stylus.

Carnegie Foundation. (n.d.). *Professional and graduate education.* Retrieved from http://www.carnegiefoundation.org/previous-work/professional-graduate-education

Carney, J. (2006). Analyzing research on teachers' electronic portfolios: What does it tell us about portfolios and methods for studying them? *Journal of Computing in Teacher Education*, *22*, 89-97.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students: A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*(3), 231-264. doi:10.1007/s11409-008-9029-x

Eitel, F., & Steiner, S. (1999). Evidence-based learning. *Medical Teacher*, *21*(5), 506-512.

Ewell, P. (2006). *Making the grade: How boards can ensure academic quality*. Washington, DC: Association of Governing Boards of Colleges and Universities.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analysis using G*Power 3.1: Tests for correlation and regression analysis. *The Psychonomic Society, 41*(4), 1149-1160.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*(3), 445-476. doi:10.1037/0022-0663.99.3.445

Gustafson, K., & Bennett, J. (2002). Promoting learner reflection: Issues and difficulties emerging from a three-year study. Mesa, AZ: Air Force Research Laboratory.

Hill-Jackson, V., & Lewis, C. W. (2010). *Transforming teacher education: What went wrong with teacher training, and how we can fix it*. Sterling, VA: Stylus.

Jenson, J. D. (2011). Promoting self-regulation and critical reflection through writing students' use of electronic portfolio. *International Journal of ePortfolio, 1*(1), 49-60.

Kember, D., McKay, J., Sinclair, K., & Wong, F. K. Y. (2008). A four-category scheme for coding and

assessing the level of reflection in written work. *Assessment & Evaluation in Higher Education, 33*(4), 369-379. doi:10.1080/02602930701293355

McIntyre, C., & Dangel, J. (2009). Teacher candidate portfolios: Routine or reflective action? *Action in Teacher Education*, *31*(2), 74-85. doi:10.1080/01626620.2009.10463519

Meier, D. (2000). *Will standards save public education?* Boston, MA: Beacon Press.

Millett, C. M., Payne, D. G., Dwyer, C. A., Stickler, L. M., & Alexiou, J. J. (2008). *A culture of evidence: An evidence-centered approach to accountability for student learning outcomes*. Princeton, NJ: Educational Testing Service. Retrieved from http://eric.ed.gov/PDFS/ED499994.pdf

Mislevy, R. J., Almond, R. G., Lukas, & J. F. (2004). *A brief introduction to evidence-centered design* (CSE Report 632). U.S. Department of Education. Los Angeles, CA: National Center for Research on Evaluation

Moss, P. A. (Ed.). (2007). Evidence and decision making [Special issue]. *Yearbook of the National Society for the Study of Education*, *106*(1).

National Council for Accreditation of Teacher Education (NCATE). (2012). *The unit standards*. Retrieved from http://www.ncate.org

Pecheone, R. L., & Stansbury, K. (1996). Connecting teacher assessment and school reform. *Elementary School Journal, 97*(2), 163-177. doi:10.1086/461860

Pellegrino, J. (2006, November). *Rethinking and redesigning curriculum, instruction, and assessment: What contemporary research and theory suggests*. National Center on Education and the Economy for the New Commission on the Skills of the American Workforce. Retrieved from http://www.skillscommission.org

Porter, A., & Smithson, J. (2001). Defining, developing, and using curriculum indicators (Report No. CPRE-RR-Ser-048). Philadelphia, PA: Consortium for Policy Research in Education. Retrieved from http://www.cpre.org/Publications/rr48.pdf

Resnick, L., Rothman, R., Slattery, J., & Vranek, J. (2004). Benchmarking and alignment of standards and testing. *Educational Assessment, 9*(1-2), 1-27. doi:10.1080/10627197.2004.9652957

Scharton, K. (2013). Reforming nurses: Historicizing the Carnegie foundation's report on educating nurses. *Nursing History Review, 21*(1), 97-108. doi:10.1891/1062-8061.21.97

Schechter, E. (20012). *Internet resources for higher education outcomes assessment*. North Carolina State University, University Planning & Analysis. Retrieved from http://www2.acs.ncsu.edu/UPA/assmt/resource.htm

Shepherd, C., & Hannafin, M. (2011). Supporting preservice teacher inquiry with electronic portfolios. *Journal of Technology and Teacher Education, 19*(2), 189-207.

Shulman, L. (1998). Teacher portfolios: A theoretical activity. In N. Lyons (Ed.), *With portfolio in hand: Validating the new teacher professionalism* (pp. 23-37). New York, NY: Teachers College Press.

Slavin, R. E. (1984). *Research methods in education: A practical guide*. Englewood Cliffs, NJ: Prentice-Hall.

Smith, S. B., Cook, C. M, Faulkner, S. A., & Peers, M. G. (2007, April). *Electronic and paper-based teaching portfolios: Student perceptions and recommendations*. Paper Presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

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), 411-433.

Urban, W. J., & Wagoner, J. L. (2009). *American education* (4th ed.). New York, NY: McGraw Hill.

Wilhelm, L., Puckett, K., Beisser, S., Wishart, W., Merideth, E., & Sivakumaran, T. (2006). Lessons learned from the implementation of electronic portfolios at three universities. *Techtrends*, *50*(4), 62-71. doi:10.1007/s11528-006-0062-9

Yao, Y., Aldrich, J., Foster, K., & Pecina, U. (2009). Preservice teachers' perceptions of an electronic portfolio as a tool for reflection and teacher certification. *Journal of Educational Research & Policy Studies*, *9*, 25-43.

Yao, Y., Thomas, M., Nickens, N., Downing, J., Burkett, R. S., & Lamson, S. (2008). Validity evidence of an electronic portfolio for preservice teachers. *Educational Measurement: Issues and Practice, 27*(1), 10-24. doi:10.1111/j.1745-3992.2008.00111.x

Zeichner, K., & Wray, S. (2001). The teaching portfolio in U.S. teacher education programs: What we know and what we need to know. *Teaching and Teacher Education, 17*(5), 613-621. doi:10.1016/S0742-051X(01)00017-8

DAVID W. DENTON is an Assistant Professor at Seattle Pacific University. He teachers graduate students in the school of education. His primary research interest includes using instructional technology to improve student learning in online and traditional classroom settings.

Acknowledgements

This research was supported in part by a grant from the Association of Independent Liberal Arts Colleges for Teacher Education.

Appendix
Writing Quality Rubric

	0 - Deficient	1 - Progressing	2 - Competent	3 - Proficient	4 - Exemplary
Artifacts	No evidence, reference, or description of artifacts appears.	Evidence from artifacts is shown without any interpretation or evaluation.	Evidence from artifacts is shown with some interpretation and evaluation, but not enough to develop a coherent analysis or synthesis.	Evidence from artifacts is shown with enough interpretation and evaluation to develop a coherent analysis or synthesis.	Evidence from artifacts is shown with enough interpretation and evaluation to develop a comprehensive analysis or synthesis.
Analysis and Evaluation in Comparison to Professional Standard	No evidence to show integration of professional standards.	Identifies connections between professional standards and coursework assignments or reflections to show an emergent understanding of teaching and learning.	Analyzes connections between professional standards and coursework assignments or reflections showing an understanding of teaching and learning.	Meaningfully evaluates and synthesizes professional standards with coursework assignments and reflections to deepen understanding of teaching and learning.	Meaningfully evaluates and synthesizes professional standards with coursework assignments and reflections along with integrating other experiences to deepen understanding of teaching and learning.

Criterion and cell descriptions were derived from "Valid Assessment of Learning in Undergraduate Education" (VALUE) rubrics, authored by the Association of American Colleges and Universities (2012).