

Online Scoring vs. Materials Scoring for Portfolio Assessments: An Exploration of Score Stability

Research Report

Hua Wei

October 2012

About Pearson

Pearson, the global leader in education and education technology, provides innovative print and digital education materials for preK through college, student information systems and learning management systems, teacher licensure testing, teacher professional development, career certification programs, and testing and assessment products that set the standard for the industry. Pearson's other primary businesses include the Financial Times Group and the Penguin Group. For more information about the Assessment & Information group of Pearson, visit <http://www.pearsonassessments.com/>.

About Pearson's Research Reports

Pearson's research report series provides preliminary dissemination of reports and articles prepared by TMRS staff, usually prior to formal publication. Pearson's publications in .pdf format may be obtained at: <http://www.pearsonassessments.com/research>.

Abstract

The study is designed to investigate whether scores assigned to portfolio submissions are comparable between materials-based scoring and online scoring conditions, and to evaluate how scorers perceive the ease of using the online scoring platform and in facilitating the scoring process. A sample of portfolio submissions from the National Board for Professional Teaching Standards (NBPTS) assessments, previously scored under the materials-based scoring condition, was selected and rescored in an online scoring platform. Comparison of score data showed that online scores were very similar to materials-based scores and that the interrater agreement statistics were also similar between the two scoring platforms. The study produced favorable results supporting the use of the online scoring system in the scoring of these portfolio submissions. Considering the advantages of online scoring and the flexibilities it offers, transitioning to online scoring is an increasingly common trend in portfolio assessments and a worthwhile effort.

Keywords: online scoring, portfolio assessment, score stability

Online Scoring vs. Materials Scoring for Portfolio Assessments:

An Exploration of Score Stability

Introduction

Portfolio assessment provides a sustained, in-depth assessment strategy for examinees to demonstrate their efforts, progress, and achievements in a given subject area (Paulson, Paulson, & Meyer, 1991). Portfolios are suitable for assessing higher-order thinking skills, problem-solving skills in real-life scenarios, and other constructs that are hard to measure with standardized tests (Miller & Legg, 1993). At the K-12 education level, portfolios are often used in statewide alternative assessment programs to evaluate the learning process and learning outcomes of students with severe cognitive disabilities. Portfolios are also used in certification/licensure testing programs for candidates to showcase their best performances in real-life workspace settings. Portfolios are effective in bringing assessment in line with instruction and can be used for either summative or formative purposes (Erickson, Niess, & Gfeller, 2000).

Traditional portfolio assessments are materials-based, requiring examinees to submit work samples, video recordings of performance, self-reflections, and other components in the form of paper copies, DVDs or CDs, or other tangible artifacts. The submission, storage, and transportation of portfolio materials require time and human and material resources. However, increased access to computers and the rapid development of electronic technology have facilitated the transition of some portfolio assessments from a materials-based format to an electronic format (Barrett, 1996; Lankes, 1995). Compared to materials-based portfolios, electronic portfolios, also called ePortfolios, contain the same types of information but are submitted and stored electronically. Electronic submission significantly reduces the cost and

efforts required of materials handling and logistics, and it facilitates the management of portfolios.

Changing the portfolio submission format presents opportunities to streamline materials handling and scoring processes. An online scoring system also capitalizes on the efficiencies offered by electronic submissions. However, changing the scoring system will impact scorer training and scoring procedures and may affect scores and proficiency level classification decisions. The stability of scores between the two scoring platforms is of utmost concern and merits investigation. There is scant research that addresses the comparability of materials-based and online scoring platforms in the context of portfolio assessments. A study by Brickley, Schwartz, and Suen (2000) compared scores assigned to materials- and web-based portfolios, but the intent of that study was to examine whether any potential score differences resulted from differences in the medium of presentation.

In this study, a sample of portfolio submissions from the National Board for Professional Teaching Standards (NBPTS) assessments, previously scored under the materials-based scoring condition, was rescored in an online scoring system with the capacity to score audio and video performance portfolios. The purpose of the study was to investigate the comparability of scores assigned under online scoring and materials-based scoring conditions, and to survey scorers' and trainers' perceptions about the ease of scoring in the two conditions. Specifically, the study addressed the following three questions:

1. Was there a significant difference in scores between the two scoring platforms?
2. Was there a difference in the interrater agreement statistics between the two scoring platforms?

3. Did scorers and trainers perceive the online scoring platform as facilitating portfolio scoring?

Method

Samples

The National Board for Professional Teaching Standards (NBPTS) offers advanced certification programs for accomplished teachers in 25 educational specialty areas. Each certificate area assessment consists of four portfolio submissions and six assessment center exercises. The portfolios are designed to capture a candidate's authentic teaching practice in real-life classroom settings. Portfolio submissions for each certificate area assessment are of different types, including one classroom-based portfolio with accompanying student work, two classroom-based portfolios that require video recordings of interactions between the candidate and his or her students, and one documented accomplishments portfolio that provides evidence of the candidate's accomplishments. Each portfolio submission is created and assembled by the candidate over a sustained period of time. Currently, all portfolio components—including samples of students' work, video recordings, documentation of accomplishments, and written commentaries—are captured on paper, DVDs, and other approved materials, and submitted using shipping courier services. Operational scoring of NBPTS portfolio submissions is then completed by trained scorers at scoring centers. The scoring process involves reviewing submitted portfolio materials, either paper copies or video segments, and assigning scores by filling out score sheets. Scores are saved in the system by scanning the score sheets.

In the pilot study, a sample of portfolio submissions was selected from an accessible pool of NBPTS portfolio submissions that had previously been scored in the operational materials-based condition, converted into electronic files through scanning and digitizing, uploaded to a

server, and rescored in the online scoring system. A total of 40 portfolio submissions were sampled from each of the four certificate areas and represented each classroom-based portfolio type. The selected portfolio submissions were originally submitted in 2010 or 2011 and were double-scored under the materials-based scoring condition. A preliminary inspection of sampled submissions verified that they were free of video issues, previously assigned a numeric score, and could be successfully converted to scoreable versions for the online scoring system. During the sampling process, cases were filtered to include only those submissions that had no resolution or backreading overrides under the materials-based scoring condition. In addition, a sampling scheme was implemented to take into account the empirical score distributions observed of any given portfolio type in the two testing years.

Online Scoring System

The online scoring system was tailored to the NBPTS scoring project. The portfolio scoring materials, including videos and supportive PDF files for any given portfolio, were properly labeled and presented in the order described in the portfolio scoring directions. The scorers were instructed to view all of the portfolio components before assigning scores on the scoring interface, just as they were instructed during scorer training for the traditional materials-based scoring.

Scorer Training and Scoring

Scorers who participated in the pilot study were recruited from the pool of scorers who participated in the 2011 operational scoring of NBPTS portfolios. A streamlined scorer training was delivered to the participants before the start of this pilot scoring project. The trainers who participated in the scoring project were all experienced trainers and attended a one-day train-the-trainer refresher meeting in advance of the scoring sessions.

For this study, each selected portfolio submission was double-scored in accordance with existing NBPTS scoring processes. The same protocols used during operational scoring, including backreading and resolution processes, were applied throughout the study. Trainers monitored scoring and scoring quality metrics, and provided backreading or resolution scores as warranted.

Scorer Survey

In the pilot study, we developed and administered a questionnaire (see appendix) to the participating scorers and trainers to survey their perceptions about the online scoring platform and to collect suggestions about how to improve and enhance the scorer training and scoring procedures. All 20 trainers and scorers who participated in the study completed the survey.

Results

Analyses were performed on the score data assigned to the sample of selected portfolio submissions under the two scoring conditions. Each portfolio submission received a first and second score under each scoring condition. Only portfolios that did not receive a backreading or resolution score under the original materials-based scoring condition were selected in the rescoring study. Some selected portfolios were assigned backreading and/or resolution scores in the online scoring system. For each selected submission, a score of record was computed by scoring condition based on the following rules. (Similar rules were applied in the computation of final scores for NBPTS score reporting.)

- If a portfolio submission receives only a first and second score, then the score of record is the average of the first and second scores.
- If a portfolio submission receives both a first and second score and backreading score(s), then the last backreading score is the score of record.

- If a portfolio submission receives a resolution score, that score is the score of record, no matter what other scores it may receive.

We computed a score of record for each selected portfolio submission by scoring condition and aggregated the scores across the 40 submissions in each certificate area. Table 1 summarizes the statistics for the aggregate scores by certificate area and for all four certificate areas combined, as well as a comparison of the statistics between the two scoring conditions.

Table 1

Comparison of Score Statistics between Materials-Based and Online Scoring

Certificate area	Scoring mode	Number of cases	Mean	St. dev	Min	Max
All	Materials scoring	160	2.655	0.696	0.750	4.000
	Online scoring	160	2.582	0.602	1.000	4.000
Early Childhood/Generalist	Materials scoring	40	2.659	0.723	1.000	4.000
	Online scoring	40	2.653	0.601	1.000	3.500
Early Childhood – Young Adulthood/Exceptional Needs Specialist	Materials scoring	40	2.641	0.696	1.000	4.000
	Online scoring	40	2.369	0.639	1.000	3.125
Early & Middle Childhood/Literacy: Reading – Language Arts	Materials scoring	40	2.681	0.678	1.125	4.000
	Online scoring	40	2.725	0.507	1.375	3.250
World Languages (Spanish)	Materials scoring	40	2.641	0.712	0.750	4.000
	Online scoring	40	2.581	0.613	1.500	4.000

Table 1 shows that the average score of record for online scoring was slightly lower than that for materials-based scoring when we looked at cases from all the selected certificate areas combined. The same is true for each certificate area, except for the area of Early & Middle Childhood/Literacy: Reading – Language Arts, in which online scoring produced a higher

average score than materials-based scoring, but the difference was negligible. The greatest difference in average score between the two scoring conditions was found in the area of Early Childhood – Young Adulthood/Exceptional Needs Specialist, where the average score differed by 0.272. The standard deviation for the score-of-record distribution was smaller under the online scoring condition than under the materials-based scoring condition, and this pattern applied to each certificate area as well as all the certificate areas combined. The range of scores, which is another index of score dispersion, was smaller for online scoring than materials scoring for the same set of portfolio submissions in each selected certificate area.

To test whether the observed difference in average score between the two scoring conditions was significant, we applied the bootstrapping method instead of a traditional t-test because the sample size in the study did not justify the use of a parametric significance test. Bootstrapping is “a general approach to statistical inference based on building a sampling distribution for a statistic by resampling from the data at hand” (Fox, 2002, p. 1). In this study, we generated a sampling distribution of differences between the two means through bootstrapping and compared the observed difference with the sampling distribution (see Efron & Tibshirani, 1993 for more details about the bootstrapping method). We used the following procedures to test the significance of the observed difference between the two means for the selected portfolios across all the selected certificate areas/portfolio types:

1. For the 160 portfolio submissions selected for rescoring in this study, combine the scores of record under the materials-based scoring condition and the scores of record under the online scoring condition. The combined pool of scores is referred to as a “pseudo population.”

2. Draw 5,000 pairs of samples (each sample includes 160 scores) from this pseudo population with replacement and then calculate the difference between the two means for each pair of samples.
3. If the difference between the two means calculated from the data falls outside the 95% confidence interval of the sampling distribution, then conclude that there is a significant difference in the average scores of the two scoring conditions. Otherwise, conclude that there is no significant difference in the average scores of the two scoring conditions.

A similar procedure was repeated to generate the bootstrapped sampling distribution of differences between the two means for each selected certificate area. Table 2 summarizes the observed difference and the bootstrapped sampling distribution of the difference for each certificate area, as well as for all the certificate areas combined.

Table 2

Observed Difference between Two Means and Bootstrapped Sampling Distribution for the Difference

Certificate area	Observed difference	Mean of sampling distribution	St. dev. of sampling distribution	95% confidence interval	
				Lower limit	Upper limit
All	0.07344	-0.00001	0.07304	-0.14317	0.14315
Early Childhood/Generalist	0.00625	-0.00008	0.14588	-0.28600	0.28585
Early Childhood – Young Adulthood/Exceptional Needs Specialist	0.27188	-0.00126	0.15174	-0.29867	0.29615
Early & Middle Childhood/Literacy: Reading – Language Arts	-0.04375	-0.00011	0.13448	-0.26369	0.26347
World Languages (Spanish)	0.05938	-0.00096	0.14786	-0.29077	0.28885

Results from the bootstrapping procedure showed that the observed difference in average score between the two scoring conditions was not significantly different from zero when we looked at the portfolio submissions across the four certificate areas. Thus, we concluded that online scoring produced scores that, on average, were not significantly different from materials-based scores. Early & Middle Childhood/Literacy: Reading – Language Arts was the only certificate area where the online mean score of record was greater than the materials-based mean score of record, but this effect was not statistically significant based on the bootstrapped samples. A nonsignificant result was also reported for each other studied certificate area. These findings suggest that scoring format did not have a significant impact on the assigned scores and that there was not a significant interaction effect between scoring format and certificate area on the assigned scores.

In addition to score differences, we looked at interrater agreement statistics to compare the scoring quality of the two scoring conditions. In the pilot study, all the selected portfolio submissions were double-scored under each scoring condition, either materials-based or online. We looked at the first and second score assigned to each portfolio submission and computed the absolute difference between the two scores. The difference score was computed for each portfolio under each scoring condition. We compared the difference scores between the two scoring conditions for the selected portfolios in each certificate area/portfolio type, as well as the four selected certificate areas combined. Table 3 presents the frequency and percentage of cases in which scores differed by a specific amount under each scoring condition. Perfect agreement between scorers results in a difference of zero score points, shown as 0.00 in the table. The column labeled “> 1.25” represents the percentage of cases that required adjudication. A case was adjudicated by a trainer if the two scores differed by more than 1.25 points.

Table 3

Comparison of Score Difference Distributions between Materials-Based and Online Scoring

Certificate area	Scoring mode	0.00	0.25	0.50	0.75	1.00	1.25	>1.25
All	Materials scoring	53 (33.1%)	25 (15.6%)	10 (6.3%)	31 (19.4%)	40 (25.0%)	1 (0.6%)	
	Online scoring	75 (46.9%)	37 (23.1%)	5 (3.1%)	16 (10.0%)	19 (11.9%)	5 (3.1%)	3 (1.9%)
Early Childhood/Generalist	Materials scoring	18 (45.0%)	2 (5.0%)	2 (5.0%)	7 (17.5%)	11 (27.5%)		
	Online scoring	22 (55.0%)	4 (10.0%)	1 (2.5%)	5 (12.5%)	7 (17.5%)	1 (2.5%)	
Early Childhood – Young Adulthood/Exceptional Needs Specialist	Materials scoring	16 (40.0%)	4 (10.0%)	2 (5.0%)	5 (12.5%)	13 (32.5%)		
	Online scoring	22 (55.0%)	9 (22.5%)	0 (0.0%)	3 (7.5%)	5 (12.5%)		1 (2.50%)
Early & Middle Childhood/Literacy: Reading – Language Arts	Materials scoring	9 (22.5%)	11 (27.5%)	1 (2.5%)	9 (22.5%)	10 (25.0%)		
	Online scoring	13 (32.5%)	13 (32.5%)	2 (5.0%)	5 (12.5%)	5 (12.5%)		2 (5.0%)
World Languages (Spanish)	Materials scoring	10 (25.0%)	8 (20.0%)	5 (12.5%)	10 (25.0%)	6 (15.0%)	1 (2.5%)	
	Online scoring	18 (45.0%)	11 (27.5%)	2 (5.0%)	3 (7.5%)	2 (5.0%)	4 (10.0%)	

As shown in Table 3, online scoring was identified with a noticeably higher rate of perfect agreement than the operational materials-based scoring. This is true for each of the four studied certificate areas. Compared to materials-based scoring, online scoring was also associated with a higher percentage of cases with difference scores less than or equal to 0.50. This seems to suggest that online scoring tends to result in a higher degree of score proximity than materials-based scoring.

However, online scoring also produced more incidents of larger disagreements (score differences equal to or greater than 1.25) than materials scoring in every studied certificate area. Cases with a score difference greater than 1.25 were routed to trainers for resolution. As mentioned above, none of the selected cases required resolution under the materials-based scoring condition. However, a few cases received discrepant first and second scores and were routed to resolution under the online scoring condition. This was later found out to be a result of the backreading functionality of the online scoring system, which worked differently from the materials-based backreading query process. Within the materials-based scoring system, portfolios with score differences greater than 1.25 were often targeted for backreading, and a backreading score was applied that eliminated the need for resolution. However, backreading in the online system did not prevent a response from being routed to resolution, so trainers and scoring supervisors typically excluded responses requiring resolution from their backreading queries to avoid scoring the responses twice—once during backreading and again during resolution. Stated another way, the higher resolution rate associated with online scoring did not necessarily mean a lower degree of scorer consistency, but rather it was a product of an online scoring functionality aimed at achieving efficiency.

A post-study survey was developed to evaluate how scorers and trainers think about the ease of scoring using the online scoring system and the usefulness of the system in facilitating the scoring process. The survey was administered online to the twenty scorers and trainers who participated in the scoring study. All participants had experience scoring portfolios under the materials-based condition. Nine of them had previously participated in online scoring projects, and six had participated in scoring projects that used other online scoring systems. Table 4 summarizes the scorers' responses to the closed-ended questions in this survey.

Table 4

Scorer Perceptions of the Online Scoring Interface and the Scoring Experience

Question	Strongly disagree	Disagree	Agree	Strongly agree
1. The flow of work within the online scoring interface is satisfactory.	0 (0%)	0 (0%)	9 (45%)	11 (55%)
2. The layout of the scoring interface is clear.	0 (0%)	0 (0%)	10 (50%)	10 (50%)
3. The online scoring interface allowed me to toggle between candidate submissions and rubrics on an as-needed basis.	0 (0%)	0 (0%)	13 (65%)	7 (35%)
4. The multiple displays offered by the two monitors allowed me to view all materials as needed.	0 (0%)	0 (0%)	4 (20%)	16 (80%)
5. The online scoring interface facilitated watching the videos.	2 (10%)	1 (5%)	5 (25%)	12 (60%)
6. The navigation controls for the video component (e.g., play, pause, jump to a point in the video by clicking the progress bar, replay, and full-screen viewing) were easy to use.	2 (10%)	1 (5%)	5 (25%)	12 (60%)
7. The video played without technical issues (i.e., pauses in playback).	2 (10%)	3 (15%)	6 (30%)	9 (45%)
8. The online scoring interface facilitated review of the written components (PDF files).	0 (0%)	0 (0%)	9 (45%)	11 (55%)
9. The training prepared me to use the online scoring interface.	0 (0%)	0 (0%)	5 (25%)	15 (75%)
10. The scoring process is efficient using the online scoring interface.	0 (0%)	0 (0%)	10 (50%)	10 (50%)

11. Overall, the online scoring interface facilitates the scoring of portfolios.	0 (0%)	0 (0%)	8 (40%)	12 (60%)
--	-----------	-----------	------------	-------------

Note. The highlighted response category for each question is the category with the largest percent of responses.

The study participants favored the setup of the online scoring interface, the quality of scorer training, and the scoring experience with the online system. Only three of the eleven questions received negative responses, and all three were related to technical details of the online scoring interface (i.e., video viewing). In addition, the scorers provided many favorable comments about online scoring. Many remarked that online scoring was an enhancement to materials-based portfolio scoring. Some scorers also commented that scoring portfolios with the online system was less time-consuming and less error-prone than scoring with the materials-based platform, because the online system could screen out problems that may cause scoring delays or errors.

Conclusions

The purpose of the NBPTS pilot study was to examine whether there were differences in assigned portfolio scores by scoring platform and to evaluate how scorers perceived the ease of use and helpfulness of the online scoring system in facilitating the scoring process. A sample of portfolio submissions that covered a range of certificate areas and portfolio types and had previously been scored in the operational materials-based condition was rescored in the online scoring system. Comparison of score data showed that online scores were very similar to materials-based scores in each of the selected certificate areas and portfolio types. The interrater agreement statistics were also similar between the two scoring conditions. Online scoring resulted in a few of the selected cases being routed to resolution. This could be a result of the

online backreading and resolution query process that differed from the operational materials-based scoring process and did not necessarily indicate more cases with discrepant scores.

Scorer survey responses showed that the participating scorers and trainers in the pilot study welcomed the online scoring system. Scorers were very positive about using the online scoring system to score NBPTS portfolio entries and were satisfied with the scoring interface. Although a few scorers mentioned that they had trouble with the video-watching features of the online scoring system, many indicated in their responses that they preferred online scoring to materials scoring. They also commented that the pre-study scorer training was adequate and effective, and it prepared them well to use the online scoring system for portfolio scoring.

Overall, this study yielded results that supported the use of the online scoring system to score NBPTS portfolios. The assigned scores were comparable between materials-based and online scoring platforms, and scorers all favored the online scoring system over materials-based scoring to score portfolios. No significant issues were discovered or reported during the study or in the study results. Although the conclusions from the study may not be directly generalizable to all portfolio assessments, which differ in substance and complexity from the NBPTS portfolio components, this study provides some guidance about the design and implementation of, as well as the expectations for a score stability study. In addition, the positive study findings address common concerns about score fluctuations due to a change in scoring platform and provide a clear indication of the equivalence of ratings between the two different scoring conditions for portfolio assessments. However, the rater-reported problems with the video-viewing features of the online scoring system need to be addressed with caution. Online scoring brings in greater efficiency to portfolio scoring, but at the same time it sets a higher demand on technology. All

the hardware- and software-related technical issues should be resolved during system transition to avoid causing problems in operational scoring.

Online scoring can improve the efficiency of the scoring process and is becoming an increasingly common trend in educational assessment scoring. With the advances in electronic technology and a growing access to computers and the Internet, ePortfolios will be incorporated more and more into educational assessments. Considering the advantages of online scoring and the flexibilities it offers, transitioning to online scoring is a worthwhile effort for portfolio assessments.

References

- Barrett, H. C. (1996). Technology-supported portfolio assessment. In R. Fogarty (Ed.), *Student portfolios: A collection of articles* (pp. 127–137). Palatine, IL: IRI/Skylight Training and Publishing, Inc.
- Brickley, D., Schwartz, A. I., & Suen, H. K. (2000). *Web-based assessment: Validation of electronic portfolios*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York, NY: Chapman and Hall.
- Erickson, D. K., Niess, M. L., & Gfeller, M. K. (2000). *Formative and summative portfolio assessment in a preservice secondary mathematics teacher education program*. Paper presented at the annual meeting of the Association for Mathematics Teacher Education, Charlotte, NC.
- Fox, J. (2002). *Bootstrapping regression models: Appendix to An R and S-PLUS companion to applied regression*. Retrieved May 24, 2012, from <http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-bootstrapping.pdf>
- Lankes, A. M. D. (1995). *Electronic portfolios: A new idea in assessment*. ERIC Digest. Syracuse, NY: ERIC Clearinghouse on Information and Technology, Syracuse University. Retrieved May 24, 2012, from <http://www.eric.ed.gov/PDFS/ED390377.pdf>. (ERIC Identifier: ED390377)
- Miller, M. D., & Legg, S. M. (1993). Alternative assessment in high stakes environment. *Educational Measurement: Issue and Practice*, 12(2), 9–15.

Paulson, F. L., Paulson, P. R., & Meyer, C. A. (1991). What makes a portfolio a portfolio?

Educational Leadership, 58(5), 60–63.

Appendix

Scorer Questionnaire

Instructions: This questionnaire is designed to gather information about the ease of scoring portfolios using an online scoring interface. Your responses will remain strictly confidential. Completing this questionnaire should take no longer than 10 minutes.

Please select the one option that best describes your response.

Have you previously participated in the scoring of NBPTS portfolios?

- Yes
- No

Have you previously trained assessors to score NBPTS portfolios?

- Yes
- No

Have you previously participated in any other scoring project that uses a Pearson online scoring system?

- Yes
- No

Please answer the following questions about THIS online portfolio scoring experience.

Please indicate your level of agreement with the following statements.

(SD = Strongly Disagree, D = Disagree, A = Agree, and SA = Strongly Agree)

Statement	SD	D	A	SA
1. The flow of work within the online scoring interface is	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

satisfactory.

2. The layout of the scoring interface is clear.
3. The online scoring interface allowed me to toggle between candidate submissions and rubrics on an as-needed basis.
4. The multiple displays offered by the two monitors allowed me to view all materials as needed.
5. The online scoring interface facilitated watching the videos.
6. The navigation controls for the video component (e.g., play, pause, jump to a point in the video by clicking the progress bar, replay, and fullscreen viewing) were easy to use.
7. The video played without technical issues (i.e., pauses in playback).
8. The online scoring interface facilitated review of the written components (PDF files).
9. The training prepared me to use the online scoring interface.
10. The scoring process is efficient using the online scoring interface.
11. Overall, the online scoring interface facilitates the scoring of portfolios.

12. Have you previously participated in a scoring project that uses a different online scoring system?

- Yes
- No

If yes, what scoring system?

13. What suggestions do you have for improving the setup of the online scoring interface?

14. What suggestions do you have for improving the training for using the online scoring interface?

15. Do you have any comments about online scoring compared to materials-based portfolio scoring?