Advances in Engineering Education

FALL 2017

You Be the Judge: When Competitions Employ an Engineering Design Rubric

GAIL LYNN GOLDBERG
Gail Goldberg Consulting
Ellicott City, MD

ABSTRACT

This article examines the use of an engineering design rubric by judges for three different student competitions—one regional, one national, and one global—to evaluate portfolios posted on the Innovation Portal, a free online resource available to students, teachers, and others engaged in STEM education across instructional levels. Judges responded to an online survey on the Engineering Design Process Portfolio Scoring Rubric (EDPPSR) following each competition. From close examination of survey responses and analysis of judges' independently assigned scores for various elements (or score categories) under which e-portfolio contents were organized, the author determines that a complex rubric can be used to evaluate competition entries both efficiently and with moderate consistency—even without more than cursory training—by those lacking familiarity with that scoring tool from an instructional context. The author identifies rewards in using rubrics like the EDPPSR for more than teaching and testing and presents some considerations regarding rubrics—not only engineering design rubrics but those intended to evaluate other processes and products common to STEM disciplines.

Key words: rubrics, engineering design, portfolio assessment, scoring

BACKGROUND

Today, decades after the term "rubric" began to enter common parlance among educators, one might well ask, "What more is left to say about rubrics that has not already been said?" Articles, book chapters, and entire volumes have been written about rubrics—what they are, what types exist, how they are created, and how and why they should be used. Many of those involved in education were first introduced to rubrics in the 1970s and 80s as scoring tools to facilitate the direct assessment of writing (Yancy, 1999). Over subsequent decades, the use of rubrics to inform evaluative judgments about student performance extended to virtually every content area, particularly as open-ended constructed response items and portfolios were added to the K-12 assessment toolkit. Ironically, the classroom use of rubrics to inform teaching and learning that was widely promoted in the 1990s and early 2000s grew, even as the use of these tools by high-stakes assessment programs diminished when the No Child Left Behind (NCLB) Act of 2001 prompted nearly all states to reduce or eliminate direct assessment. Rubrics were—and continue to be—celebrated as a useful and effective means of making learning targets clear to students, while providing teachers with an efficient and objective means of conducting evaluations of student work. Their value to students and teachers of engineering design is evidenced by the many examples of rubrics addressing this aspect of engineering available online. One single site alone, http://www.rcampus.com/indexrubric.cfm, turns up hundreds under the key phrase "engineering design process".

The belief in the capacity of a rubric to provide a common language and set of expectations has recently informed a promising new practice—the shared use of an engineering design process rubric by students engaged in innovative projects and practicing professionals and college faculty who serve as judges of design and problem solving competitions. Over the past few years, an increasing number of competitions—state, regional, national and international—have turned to the Engineering Design Process Portfolio Scoring Rubric (EDPPSR) as both the framework for student projects and the criteria by which a diverse community of judges including academics and practicing professionals have evaluated student projects and determined those to whom to bestow honors and awards. (Consult https://www.innovationportal.org/resource/samples to access the full suite of rubrics that constitutes the EDPPSR, along with scored sample entries; see http://teams.mspnet.org/media/data/EDPPSR.pdf?media_000000008449.pdf for rubric only). Anyone who has witnessed or actively engaged in an effort to adopt a particular rubric or suite of rubrics is likely to recall the struggles that endeavor involved. Therefore, the experiences thus far of competition judges using the same rubric that was designed to inform instruction, coupled with examination of their scores, offer some new insights into this decades-old approach to classroom and large-scale assessment.

THE EDPPSR: ORIGINS, INTENDED USES, AND NEW OPPORTUNITIES FOR USE

It was mutual interest in the idea of a rubric that could be used to evaluate evidence of the engineering design process that first led to collaboration between Project Lead The Way (PLTW), a non-profit organization devoted to the design and implementation of cutting-edge computer science, engineering and biomedical science curricula, the Kern Family Foundation, and a group of engineering educators seeking a pathway to college admissions and advanced placement (see Abts, 2011; Groves et. al., 2012; Groves et. al, 2014; Robelen, 2013). In 2010, a draft rubric originally developed by Mark Schroll, former Director of Strategic Initiatives for PLTW, as the cornerstone of an online e-portfolio he envisioned was scrutinized by a team of educators and engineering practitioners under the leadership of Dr. Leigh Abts, University of Maryland, with funding support from the National Science Foundation (NSF) and the Kern Family Foundation. Afterward, the principal project investigators sought the aid of the author, an assessment specialist with expertise in rubric design and development, to revise that draft. Subsequent piloting and further revision in 2011 yielded the version of the EDPPSR in use today (Goldberg, 2014). It was recognized almost immediately as a strategy for assessing the engineering design process that might be applied not only by teachers but also by "external reviewers with a vested interest in the design solution." (Householder & Hailey, 2012, p. 31). Concurrently, that version became the framework for the Innovation Portal, a free online resource available to students, teachers, and others engaged in STEM education and the focus of an NSF PRIME grant to examine the validity and reliability of the EDPPSR (National Science Foundation, 2011).

In the years since the EDPPSR evolved, exposure to the rubric via the Innovation Portal has greatly increased. Not only is the rubric being used to assess the portfolios submitted by thousands of students for PLTW's Engineering Design and Development (EDD) capstone course, but students in other instructional contexts as well—from middle school through college—are using the rubric-based framework to organize entries and artifacts that demonstrate their engagement in the engineering design process. (Live, to-the-second data on users of the framework is displayed on https://www.innovationportal.org/statistics). Furthermore, a key objective of those who created the Innovation Portal—to induce students to create and share their best work by using that tool to host design competitions with real rewards—is being addressed through a growing array of organizations and industry sponsors (www.innovationportal.org/faq; also see https://www.innovationportal.org/opportunities for current and past competition opportunities). The capacity and value of the EDPPSR to serve as a framework for work that may have been created outside of a specific curriculum and evaluated by judges who for the most part have not previously used, or even been exposed to, the rubric was the focus of this study.

A TALE OF THREE COMPETITIONS: DATA SOURCES AND METHODOLOGY

Among various competitions hosted via the Innovation Portal in 2015-2016 were—in chronological order—the KEEN EMpwr Competition, the Conrad Spirit of Innovation Challenge, and the Kansas

Table 1: Elements of the Engineering Design Process Portfolio Scoring Rubric (EDPPSR)

Component I: Presenting and Justifying a Problem and Solution Requirements

- Element A: Presentation and justification of the problem
- Element B: Documentation and analysis of prior solution attempts
- Element C: Presentation and justification of solution design requirements

Component II: Generating and Defending an Original Solution

- Element D: Design concept generation, analysis, and selection
- Element E: Application of STEM principles and practices
- Element F: Consideration of design viability

Component III: Constructing and Testing a Prototype

- Element G: Construction of a testable prototype
- Element H: Prototype testing and data collection plan
- Element I: Testing, data collection and analysis

Component IV: Evaluation, Reflection, and Recommendations

- Element J: Documentation of external evaluation
- Element K: Reflection on the design project
- Element L: Presentation of designer's recommendations

Note: The suite of rubrics that comprise the EDPPSR include these twelve, which provide the framework for the Innovation Portal and entries for which are scored based on the contents of that section of the portfolio. In addition, there are two Elements under Component V (Documenting and Presenting the Project) that may be applied holistically to evidence in the portfolio as a whole:

- Element M: Presentation of the project portfolio
- Element N: Writing like an engineer

City STEM Alliance Competitions for Engineering Design and Biomedical Research and Design—each with a different scope, stakes, and participant target. These three competitions utilized the EDPPSR somewhat differently but shared the expectation that students would post evidence of their engagement in the design process in an e-portfolio organized according to the rubric elements describing aspects of that process (see Table 1 for EDPPSR Elements and Components of the engineering design process under which they are organized). The portfolios would then be scored using the criteria articulated in the rubric.

The author sought and received from each competition's governing body assistance in obtaining feedback from judges about their training on, and use of, the EDPPSR. Following each competition, all judges received from their competition administrator an emailed invitation to voluntarily respond to an online questionnaire about the EDPPSR intended to inform research on the rubric (see Appendix A or http://goo.gl/forms/qnGDj6sSmW). They were not apprised of the survey before the completion of judging. The administrator's email and the introductory matter preceding the survey questions made clear that judges were under no obligation to participate. They were informed of the purpose of this data collection and the likely approximate time needed to respond, and were assured of confidentiality of responses. Response rates from the three sets of judges varied from 27-41%. The author did a content analysis of the judges' survey feedback, leading to quantification of response types.

The second data source was the Element scores assigned by every judge in the three competitions. All users registering on the Innovation Portal are directed to a privacy policy and terms of use agreement (including sharing of score data with third parties), which they must accept in order to proceed. Prior to the competition opening dates, the author obtained a Memorandum of Understanding with PLTW allowing for provision of competition scores, and was also granted access by each sponsoring organization. Once the current Innovation Portal administrator collected and provided scoring data (after all identifying information had been removed), the author analyzed these data for patterns in rater agreement.

In every instance, judges had received only a very brief training of approximately one hour on how to navigate the Innovation Portal and enter scores online, and they were shown how to access the rubric and a set of scored and annotated sample entries for each element, across score points. There was no formal scoring training such as is customary in operational scoring for student assessments—no guided review of anchors and training samples for example. Therefore, conventional formulas for acceptable interrater agreement (interrater reliability) were regarded as inappropriate for this study. Instead, discrepant scores were flagged following guidelines typically used in preliminary rangefinding—an assessment activity in which a team of individuals with content expertise but often little/no experience with a given rubric apply those evaluative criteria independently to samples of student work before discussing and trying to reach consensus on scores for samples to be later used in training and monitoring of raters. (For a description of a common approach to rangefinding, see Maryland State Department of Education, 2008.) During rangefinding, it is not uncommon prior to discussion to see a three point spread on a six-point (e.g., 0-5) scale. Since the judges for these competitions had little/no prior experience with the EDPPSR, no systematic scoring training beforehand, and no opportunity to discuss their scores, their experience was deemed analogous to that of rangefinders first embarking upon that activity; therefore, for the purposes of this study, only score distributions outside of any three adjacent scores (e.g., 1-2-3 or 3-4-5) were considered instances of discrepancy.

All available score data for the three competitions (three entire populations of data) were examined in the context of the convenience sample of judges' feedback after each competition on the ease of use of the rubric and problems they identified.

KEEN EMpwr Competition

Competition Background

In the late fall of 2015, undergraduate students enrolled in any of the twenty-two colleges or universities that were part of the Kern Entrepreneurial Engineering Network at that time (see http://engineeringunleashed.com/keen/partners) had the opportunity to submit a design project portfolio for the KEEN EMpwr Competition. Thor Misko, Program Director and Team Leader of the Kern Family Foundation's Entrepreneurial Engineering Program and self-identified "big fan of the rubric and Innovation Portal" organized the pilot competition (Misko, December 4, 2015). For the purposes of this competition, portfolios addressing EDPPSR Elements A through D and F were to be posted to the Innovation Portal by teams of at least two, but no more than five students. The decision to focus on the "front-end" of the design process was based on anecdotal feedback from students, faculty sponsors, and judges about the alignment of the rubric with what was going on in their classrooms. Educators and practicing professionals were recruited to score the entries using the criteria for these elements of the engineering design process. Prizes to the winning team included expense-free travel to, and participation at, the KEEN winter conference, monetary awards to the winning and placing teams' college or university to support KEEN-related activities and/or programming, and—perhaps most appealingly to students—an early-stage Market and Feasibility Assessment by the Wisconsin Innovation Service Center.

Students were provided links to instructional videos on getting started using the Innovation Portal to create a portfolio and on building the portfolio by adding text, images, video, and other artifacts. Students also learned how to consult the rubric and the scored and annotated examples on the Innovation Portal to help guide their work. For nearly all student participants and judges alike, the Innovation Portal and the EDPPSR were new.

Survey Feedback

Eleven teams submitted portfolios that then were each scored by five members from the pool of twenty-two judges. Of those who participated as judges in this pilot competition, six (27%) responded to the survey on the EDPPSR: five indicated an academic affiliation and one a corporate affiliation. The rather low response rate may have been due to the fact that judging closely followed the submission deadline of November 13th, but judges did not receive the request to participate in the survey until early January 2016 (with a reminder in mid-February).

Four of the judges who responded indicated they were not familiar beforehand with the EDPPSR, while two were familiar with it—one from another competition (the Kansas City STEM Alliance Competition, described elsewhere in this article) and another from classroom use. Most cited a webinar offered by PLTW personnel as the means of training or preparation, but one mentioned that she and an experienced judge "walked through one of the submissions" together. No one indicated spending more than about 30 additional minutes in preparation beyond the hour-long webinar.

All six judges indicated that they found the process of becoming familiar with the EDPPSR straightforward and easy. The only rubric element cited as challenging to score (and then, by only one judge) was Element D, but the explanation provided made clear that the challenge was created by the fact that student submissions were at very different stages of development. All the judges cited the scored and annotated examples available online as being particularly helpful to them, when using the EDPPSR, and one indicated as helpful the specific criteria in the rubric itself.

Given the opportunity to suggest how the process of using the EDPPSR could be made easier and/or more effective, none of the respondents offered any suggestions. Instead, they focused on features of the competition itself, suggesting that it was unfair to judge students the same whether they submitted evidence of existing or new projects and whether projects were large or small in scope. One respondent felt that process would become easier simply by using the rubric more often, and noted, "My colleague an I are adopting both the Innovation Portal and the EDPPSR in our sophomore studio design course"—a reversal of the more typical route from classroom practice to competition.

With the small response rate, it is impossible to know how other KEEN EMpwr judges regarded their experience using the EDPPSR. One respondent's comments, however, pointed to what may have been a key weakness in the pilot competition. That judge noted, "I think there's a certain threshold of project depth and duration where the Portal is highly beneficial—like semester long projects, senior design, etc." Given the narrow window for submissions that captured only a few weeks' work, use of both the Innovation Portal and the EDPPSR may have been "overkill." Nevertheless, the pilot served to introduce partner universities to the EDPPSR.

KEEN EMpwr Scores

Instances of discrepancies among the assigned scores support the suggestion by one survey respondent that Element D was the most challenging to score, with 6 of 11 portfolios receiving discrepant scores in this category (in one case resulting in a six-point spread from 0-5). All but two portfolios (one mid-range and one strong) received at least one discrepant element score, with three receiving three discrepant element scores out of the five assigned for each portfolio. However, in each of these instances, there was evidence to suggest that one judge was an "outlier," assigning scores for all elements higher than, and out of range of, the scores assigned independently by the other four judges scoring that portfolio.

When scores were examined by individual element, rather than across a whole portfolio, outliers appeared at both high and low ends. Although overall, 20 of 55 possible entries for Elements A-D and F received discrepant scores from the five judges scoring each portfolio entry, that number was reduced to 4 of 55 if the high or low outlying element score was disregarded—an accepted practice in other contexts such as athletic competitions and grant reviews. For the purposes of this competition, organizers determined winners by setting a cut score and awarded a prize to the team whose cumulative score was highest after dropping high and low judges' scores for each element.

Conrad Spirit of Innovation Challenge

Competition Background

The Conrad Spirit of Innovation Challenge is one centerpiece of the Conrad Foundation established in 2006 by Nancy Conrad in honor of her late husband, astronaut and entrepreneur Charles "Pete" Conrad Jr. This annual competition, conducted for the ninth year in 2015-16, seeks to encourage high school students (ages 13-18) around the globe to demonstrate innovation and entrepreneurship by creating and presenting design solutions in the areas of Aerospace and Aviation (AA), Cyber-technology and Security (CS), Energy and Environment (EE), and Health and Nutrition (HN) (www.conradchallenge.org). Partners, including but not limited to NASA, the Kennedy Space Center, the Wisconsin Innovation Service Center, and the Milken Family Foundation, not only provide resources but also contribute to the pool of judges representing business, industry and academia. In 2014, seeking an efficient and effective way to conduct judging of student submissions, the Conrad Foundation first partnered with Project Lead The Way to utilize the Innovation Portal as the vehicle for documentation and review of team projects—including the use of the EDPPSR as the framework for portfolios and the criteria by which they would be evaluated. The competition announcement included information on the Innovation Portal and the rubric by which projects will be judged. Once registered to compete, students could open an account and learn about both portal and rubric through an online tutorial.

For the 2015-16 competitions, semi-finalists were selected based on a preliminary "Investor Pitch" of their project accompanied by a brief abstract, and invited to develop their concept in a project portfolio covering Elements A through F. Top-scoring teams were then asked to complete additional Elements G through I. All those who participated in this second round of the semi-final were invited to attend the Innovation Summit at the Kennedy Space Center Visitor Complex as Conrad Diplomats and have the opportunity to be named Pete Conrad Scholars.

Competition judges were volunteers who committed to participating in an hour-long training webinar primarily focused on navigating the Innovation Portal and introducing them to the EDPPSR. They were encouraged to review the scored and annotated sample entries available on the Innovation Portal, both in preparation for and during scoring. Every portfolio submitted was scored independently online by at least three judges, and judges typically scored 4 portfolios.

Survey Feedback

Between late February and mid-March, 2016, thirty-four Conrad Spirit of Innovation judges (out of 106 who scored at least one portfolio) responded to the request to complete the online survey on the EDPPSR after scoring Elements A-F (semi-final round Part I). Based on the total number of judges, the response rate was 32%. Among the twenty respondents who indicated an affiliation, ten identified an academic institution, five a corporate entity, three a medical institution, one a professional society, and one the Conrad Foundation. By over two to one, respondents indicated that they had no previous familiarity with the rubric (No: 24; Yes: 10) and in fact only four had actually used the rubric previously (all in the context of judging a previous Conrad competition). The remaining six respondents who indicated prior familiarity cited the training session and/or self-study of training materials as the context. One respondent who indicated "no" elaborated by noting "but very helpful and will likely use it in my classes."

The majority of respondents (11/34) identified as the training method and time the one hour mandatory training session/webinar, and six cited the online training but did not indicate time. Another ten respondents indicated training by webinar and reading examples independently; of these, five indicated training took one-two hours and five did not indicate time required. The remaining six responses mentioned time but not method or described additional training input besides the webinar (e.g., previous experience, follow-up emails).

Queried on how the process of becoming familiar with the EDPPSR might be made easier and/or more effective, among the thirty respondents who answered this question, fifteen indicated that the process worked well ("very user friendly, " "pretty simple, " quite good," "easy and straightforward"). It is worth noting that the EDPPSR was new to eleven of these fifteen respondents, while four had used the rubric previously when participated in judging an earlier competition. Suggestions for improvement included, but were not limited to providing training on one good example, so that judges could try out for themselves; conducting a guided review of sample responses; and providing access to rubric and examples in advance of the webinar.

Asked which, if any, of the rubric elements judges found more challenging to score and why (regardless of how easy or difficult it appeared for students to address those elements), a substantial number of respondents did not identify any elements that they found challenging to score: of these, thirteen indicated "none" and/or elaborated upon their response (e.g., "no troubles scoring any of them"; "they were all about the same. I used the drop down menu to look at the rubric as I was scoring to be sure that I was doing it uniformly and fairly"), five left this question blank, and one respondent regarded the elements as equally difficult to score. Of those with no issues, ten were new judges.

Among those judges who identified particular elements they found challenging to score, two elements were noted, and in some instances elaborated upon, by more than one or two judges:

Element E (5 responses) and Element F (6 responses). At least some of the judges who identified challenging elements did—in spite of instructions—focus instead on the difficulty of those elements for students (e.g., noting that "submissions I reviewed did not have a good grasp on what a STEM principle/concept was" or students' failure to demonstrate that where needed, experts have reviewed the proposed design solution). The difficulty that a few judges cited related to Elements C and D did not relate to the rubric itself but to students' occasional disregard for distribution of evidence across different sections of their portfolio that made scoring more difficult.

Conrad Spirit of Innovation judges most often cited as particularly helpful when using the EDPPSR to make score decisions the access to scored and annotated examples of student portfolio entries (cited by 15 judges, 12 of whom had no previous experience with the EDPPSR). Ten judges cited the actual scoring guidelines articulated in each element rubric. Others among the judges focused on the Innovation Portal rather than the rubric, and noted such features of the platform as immediate access to rubrics, examples, etc. (e.g., "pull up criteria while scoring") and its ease of use (8), and availability of PDF versions of the rubric and reflective questions (2).

Although some of the respondents had no suggestions to make scoring with the EDPPSR easier or more efficient, with several commenting that it was straightforward and "perfect as is," others had suggestions for "next steps" or additional resources might be helpful. These included streamlining key differences between score points for each element; stressing those features that are "a must" for a given score; and developing and providing for student teams some suggestions on how to post their portfolio entries/artifacts on the Innovation Portal so submissions will align better with the rubric and thus be easier to evaluate. Virtually all of the additional resources suggested were ones that would be easy to implement in time for future competitions. Meanwhile, the generally favorable response to use of the EDPPSR by Conrad Spirit of Innovation Challenge judges, particularly among individuals who had never used the rubric before, suggests that it was regarded as a tool that could continue to be a useful for judging students' design projects.

Conrad Spirit of Innovation Challenge Scores

Instances of discrepancies in the assigned scores suggest that Elements D and E were the most challenging for these judges to score (with 29 of 108 portfolios receiving discrepant scores for these elements), followed closely by Element F (with 28 of the 108 portfolios receiving discrepant scores). Fifty portfolios were free of any discrepant scores (spreads wider than adjacent scores). An additional twenty-one portfolios were assigned discrepant scores for only one of the first six elements, and an additional nine were assigned discrepant scores for only two elements. At the other end, in terms of agreement, only seven portfolios received 5 or 6 discrepant scores, suggesting serious issues in applying the criteria to those portfolios.

When examined by category, the greatest number of portfolios with no discrepant element scores was among those submitted to the CS category (64%). Portfolios submitted to other categories were somewhat more challenging to score, with AA 47% discrepancy-free, HN 42% discrepancy-free, and EE 41% discrepancy free. There appears a parallel when considering discrepancies across elements; the lowest percent of discrepant scores were those assigned to CS portfolios (14% discrepancies), followed by (in increasing percent of discrepancies) AA (23%), HN (27%) and EE (29%). Given that CS judges were assigned more portfolios (7-8) than judges for the other categories (who generally scored 3 or 4), it is possible that their greater experience relative to the other judges contributed to better agreement.

It is worth noting that in judging for the Conrad Spirit of Innovation Challenge, out of the many scores from so many team members there were only five instances in which the nature and scope of discrepant scores suggested that one judge on a given team was consistently an "outlier" (assigning scores for all or nearly all elements considerably higher or lower than, and out of the non-discrepant range of, the scores assigned independently by the other judges scoring that portfolio). When element scores were examined, outliers appeared at both high and low ends. Although overall, there were 150 instances in which there was a discrepancy across judging team members' element scores, that number was reduced to only 56 out of 648 (around 9%) if the high or low outlying score was disregarded.

Kansas City Stem Alliance Competitions

Competition Background

The 2016 Kansas City Engineering Design and Biomedical Research Design Competitions, sponsored by the KC STEM Alliance, provided the opportunity for students in the Kansas City region to receive recognition and scholarship funds, along with valuable mentor input on the market feasibility of their ideas. According to the competition notice, the goal of the engineering design competition was to "highlight the importance of engineering design and problem solving in STEM education" while the goal of the biomedical research design competition was to "recognize and reward outstanding student achievement in the implementation of the scientific research method or design process." Eligible students, who had the option to submit an individual portfolio or compete as part of a team, included those enrolled in Project Lead The Way's Engineering Design and Development course (EDD) or Biomedical Innovations course (BI). The expectation was that portfolios would be built around students' capstone projects for EDD or BI. The competition was organized as an incentive for students to stay "on track" with their long-term projects and the calendar coincided with course requirements and timeline.

For the KC STEM Alliance, this was the second year in which the EDPPSR was applied as scoring criteria, but only the first year in which BI students were also included. This seemed reasonable since from its inception the Innovation Portal was envisioned as adaptable beyond engineering applications to design and problem solving projects across disciplines (Carr, 2012). In the recruitment and selection of judges—all practicing professionals—as well as their eventual scoring assignment, care was taken to match expertise to the portfolio category. The notice of competition provided links to videos on creating a portfolio and accessing both the EDPPSR and scored sample entries that could serve as models. Students, teachers and judges were informed that for the purposes of the BI competition, the word "prototype" should be replaced by "experiment" when addressing all elements (G, H, and I) in Component III of the rubric, and no scores would be assigned for Element G in BI portfolios.

Survey Feedback

Of eighty-three judges (47 engineering and 36 biomedical) who scored at least one portfolio, thirty-four judges (41%) responded to the survey. Eleven respondents gave no affiliation; of the remainder, nine indicated academic affiliations, eleven professional, and three medical. Although this year's judges included eight who had scored the competition previously, all but one respondent indicated a lack of familiarity with the rubric before the competition. While the majority of respondents cited the webinar alone as their preparation for judging, nine also mentioned other activities such as reviewing emailed instructions, studying a printed copy of the rubric, reviewing the annotated samples, and reading through submissions and doing "practice scoring" before embarking on the real thing.

Nearly half of the respondents offered no suggestions for making the process of becoming familiar with the EDPPSR easier or more effective, commenting, for example, that the rubric is "good the way it is," and "just need practice with it to be prepared." Two repeated themes among comments were that it would have been helpful to have scored one common portfolio, with established scores, for practice—what one judge called a "quality assurance session" and another, "calibration"—and to provide among the scored and annotated examples some that were specific to biomedical sciences.

Among KC judges, Elements E and F were most often cited as ones that were particularly challenging to score. Judges' explanations for Element E included observations that students did not seem to understand who might constitute experts who could verify application of STEM principles and practices in the proposed design solution and a lack of understanding by students and judges alike of what was meant by application of "STEM principles and practices" (another way of saying that work was data-driven and provided evidence of the math and science that make the design viable). Element F was challenging to some judges who did not understand the meaning of "extra-functional considerations," although this could have been resolved by clicking on a reference page link in the rubric sidebar. There, the term—which refers to all the factors like cost, ease of distribution, capacity for mass manufacturing that impede something going from the bench or lab into the hands of those envisioned as users—is clarified. It is worth noting that nearly always, these comments and concerns came from judges of the biomedical competition.

As with judges from the other competitions, the KEEN EMpwr and The Conrad Spirit of Innovation Challenge, KC judges often gave off-topic responses when questioned as to which elements were most challenging to score, regardless of students' facility with them. Among those responses were ones that addressed the challenge to students (e.g., "Based on rubric most students didn't score well, though some were very good quality. This was difficult to reconcile;" "the first element [A] because students often fell short compared to what the rubric called for"). Other comments dealt more generally with the challenge of rubric-based scoring or the difficulty of applying engineering design criteria to BI entries ("I think the challenge for me as a first time user judging Biomed was ensuring my own user consistency"; "It was challenging to judge something that had nothing to do with what was asked to be done as in the rubric. There were no standards in these occasions and I had to make a judgmental call but I don't think that is the best way to evaluate").

Over half of the respondents identified the scored and annotated sample entries as being particularly helpful when using the EDPPSR ("the examples and rationales were fantastic"), although two noted that they were "not really helpful for biomed projects." The rubric itself was praised by eight judges who cited the clear progression from score point to score point and the level of detail that informed score decisions. Among comments from KC judges were some of the most detailed and thoughtful ones about the decision-making process expressed by judges across the three competitions. For example, one cited "practice" as something that was particularly helpful when using the EDPPSR, noting, "The fact that I found the second and third submissions were easier because I had learned what I was looking at one the first one." Another judge responded:

> I found it helpful to me to do my assessments one day including pasting the text for the rating I had given each element into the comments field. Then I let them "sit" overnight and came back the next day to revisit them, replaced the copied text with my own comments, finalized and submitted them.

The generally positive outlook towards the rubric among the KC judges was mirrored in responses to the subsequent survey question about how the process of using the EDPPSR to evaluate portfolios might be made easier and/or more effective. Most frequently, judges indicated that the rubric was okay or easy and effective as is (8 responses). Three suggested streamlining the rubric. Among responses to this question were two contradictory ones. One judge, who was among those who suggested that the rubric be shorter, said, "It was definitely not strongly suited to evaluating biomedical science projects that did not necessarily include design elements." Meanwhile, another noted, "Liked rubric, found fairly easy to use on Biomed projects." KC judges more consistently than those in the two other competitions entered feedback in the survey comments field when scoring. The abundance of comments led one judge to suggest, "Review judge's comments to see if relevant information could be given as feedback to the instructors working with the students-incorporate into coaching for improvement. Also, feedback to the judges-were the scores they assigned outliers?"

Although KC judges sometimes went off-topic in their responses to the question about improving the process of using the rubric, off-topic responses accounted for far fewer than was the case among judges of the earlier two competitions. One judge's comment alluded to students not following the rubric, and another raised questions about portfolios that were incomplete and may not have represented participants' best work. Rather than addressing the question, a few respondents used this response field to reiterate two common themes in KC judges' feedback: creating a competition-specific rubric ("Design the rubric for engineering and for biomedical;" "Some of the engineering criteria were not applicable to the biomedical students and thus they received lower scores based on the rubric because of missing information") and building practice on a sample portfolio into training ("I would suggest an example project portfolio that each judge could go through. Maybe once we submit our findings have some sort of way to compare with others. Feedback is priceless for judges too!").

Although the EDPPSR was new to the vast majority of KC judges, when asked how this most recent experience using the EDPPSR compared to any earlier experience(s), one of the three experienced users of the rubric who responded to the survey simply said, Excellent." The last survey question invited additional feedback and elicited a range of comments, including:

- I generally like the approach and believe it offers an opportunity for fair evaluation for each participant.
- Based on rubric criteria and examples I feel I harshly graded students/ I hope someone is spot-checking grading results for over grading.
- Encouraging young student entering science and providing wonderful platform.

Once again, however, a few judges used this response field to express concerns about using the EDPPSR for "biomedicine/lab science-based projects"; one judge among these elaborated by noting, "It did not fit very well and the actual projects I scored were universally woefully inadequate based on the rubric."

Kansas City Stem Alliance Competitions Scores

While feedback from competition judges suggests substantive issues with applying the EDPPSR to biomedical portfolios, the difference between discrepant scores assigned to portfolios from each category is not as dramatic as might be expected based on that feedback. Of element scores assigned to engineering design portfolios, 22% of scores were discrepant by the definition established for this paper, while for the biomedical engineering portfolios, 29% of the assigned scores were discrepant.

Instances of discrepancies in assigned scores suggest that among judges of the engineering design competition, Elements C and E were the most challenging to score, with eighteen of fifty-four portfolios receiving discrepant scores for these elements, followed closely by Element F with sixteen of the fifty-four portfolios receiving discrepant scores. For the biomedical research and design portfolios, based on number of discrepancies noted, Elements C and K were the most challenging to score. Collectively, across both competitions, Elements B and L generated the fewest number of discrepancies: seventeen out of ninety-seven decisions possible for each element.

Fifteen of the engineering design portfolios were free of any discrepant scores. An additional seventeen portfolios were assigned discrepant scores for only one or two of twelve elements. At the other end, in terms of performance, eight portfolios received six or more discrepant scores (15%), suggesting serious issues in applying the criteria to those portfolios. Among the biomedical portfolios, only five were completely free of discrepant scores and an additional nineteen were assigned discrepant scores for only one or two of eleven elements (with Element G not included for that competition). Again, eight portfolios from this category received six or more discrepant scores (19%). Comparison of the percentages of portfolios with a majority of discrepant score decisions by judges reinforces the idea that in spite of judges' feedback, the difference in rater consistency is not far different between one competition and the other.

To ensure that some score decisions were not identified as discrepant based on the assignment of a 0 score, which was used dually to signify the lowest score point based on features of a given portfolio entry as well as a missing entry (there being no condition code for blank or empty sections of the portfolio), notes for each 0 score were examined. Most often, the judge elaborated on the deficiencies of the entry (using the language of the rubric to explain why that score was assigned) rather than indicating "no content" in that section. Only true 0 scores were included in the determination of discrepant decisions between/among judges.

Bearing in mind that discrepancies might result from some judges' score decisions being "outliers" (considerably higher or lower than, and out of the non-discrepant range of, the scores assigned independently by the other judges scoring the same portfolio), instances when the score of one of three or four judges (or an identical score from two of five judges) differed by more than one point from the next nearest score were flagged. When outliers at either high or low ends were dropped, the total number of discrepant scores in engineering design dropped from 144 to 32, and in biomedical engineering from 137 to 36 (again, fairly comparable results from both competitions).

LOOKING AHEAD TO FUTURE COMPETITIONS

Each of these three competitions that had utilized the EDPPSR as criteria for determining commendations and awards had somewhat different plans as they looked ahead to the 2016-17 academic year and beyond.

KEEN EMpwr competition organizer Thor Misko had seen value in introducing partner universities to the EDPPSR so they might consider if and how they might leverage the tool in their classrooms and extracurricular activities. The prospect, in so doing, of creating a common language and common set of expectations, unlike the case of some other college-level engineering competitions in which each institution applied their own standards, held great appeal to him; but he recognized, post-pilot, that change would happen "one faculty member at a time," only as the benefits of using a common metric became apparent. For post-secondary students, for whom admission or placement was no longer an issue, he identified a different driver, suggesting that "if industry says they'll start to recognize [the criteria articulated in the EDPPSR] they'll pay attention to it." (Misko, December 16, 2016). Rather than hosting another KEEN EMpwr Competition in the 2016-2017 school year, the Kern Family Foundation sponsored the participation of forty-four students from twelve KEEN Partner Institutions in the University Innovations Fellows Program (http://universityinnovationfellows.org), with selection of participants made by those Partner Institutions. While the use of the EDPPSR to evaluate a very time-constrained project may have been like using a sledgehammer to crack a nut, the pilot did have the effect of introducing the rubric to a wider audience—and one that extends beyond the PLTW community—and at least some of those exposed to the rubric through the competition expressed their intention to incorporate the EDPPSR in classroom contexts in the future.

The Conrad Spirit of Innovation Challenge engaged more judges with experience using the EDPPSR than either of the other two competitions, and evidence suggests that the comfort level with the rubric grew and would continue to grow were use of the EDPPSR to continue. In the immediate aftermath of the competition, tentative plans were to continue to use the Innovation Portal as an evaluative tool and to explore the possibility of expanding resources by adding evaluative criteria and annotated samples of student work for entrepreneurship to the Innovation Portal. Since that time, however, the competition organizers and sponsors elected to switch for 2017 to a different approach to evaluation of submissions that focuses more on "the entrepreneurial elements so central to [the] competition." (Conrad, March 16, 2017).

Immediately following the 2016-17 competition, The KC STEM Alliance planned to use the EDPPSR again. However, in response to feedback she received from students, teachers, and judges before, during, and after scoring the 2015-16 competitions, Ann Zimmerman, Industry Coordinator for the Alliance, envisioned a number of improvements. Recognizing that the quality of submissions improved from the previous year, and attributing this in part to greater familiarity with the Innovation Portal and the rubric, she addressed this topic in advance of the next year's Engineering competition. In 2015-16, although all students had access to the rubric early on, EDD students based projects from their inception on the rubric while BI students completed a more traditional experimental design process and then "retrofitted" those projects to align with the EDPPSR. Zimmerman anticipated that more—and earlier—classroom conversation with BI students about rubric language and criteria would lead to improvement in the quality of their submissions. For the 2016-17 competition, this included making minor changes in the language of the rubric and a "reflective questions" resource to include reference to experimental design projects. Since competition-specific changes could not be made to the EDPPSR as it appears on the Innovation Portal, plans were to modify the rubric "offline" so that its application to experimental design is clearer. Based on questions about the appropriateness of the EDPPSR for the BI competition, the KC STEM Alliance will be exploring other possibilities as well for future years. (Zimmerman, May 2016; July 2016; March 2017).

WHAT'S THE VERDICT?

Among the various scoring schema that have been and, in some cases, continue to be used today both in the classroom and to evaluate competition entries in STEM disciplines, a variety of formats are evident. Some scoring tools are designed as activity-specific checklists on which points are assigned for discrete observable characteristics or components. (See, for example, http://rocketcontest.org/wp-content/uploads/2016/09/Engineering-Notebook-Rubric.pdf and http://nfpahub.com/wp-content/uploads/sites/4/2016/05/Judges-Rubric-2017.pdf). Such systems have their place when the goal is to simply rate and rank submissions that address the same or very similar task. Other scoring tools identify traits and provide descriptors for each under a set of performance categories like "beginning," "developing," "accomplished," and "exemplary". (See for example http://www.roboticseducation.org/documents/2014/10/vex-iq-design-award-rubric.pdf). However, sometimes these descriptors are quite cursory or mix the "apples and oranges" of process (identifying a problem, reviewing existing solutions) and product (delivery of presentation). Still others among the plethora of STEM rubrics one can access today expect users to assign scores for abstract traits like originality and creativity. Without some working definition and examples (so that those doing the judging will clearly "know it when they see it"), consistency among judgments about such elusive traits can be especially challenging to achieve.

Another feature of many competitions' scoring criteria is that they are designed only for a specific occasion—the evaluation of entries in order to determine those worthy of awards and recognition—or a specific product such as a submersible or a drone. While the best competition criteria mirror academic and professional standards and expectations for the disciplines to which the entry is most closely related (whether engineering, biomedical research, or some other), the "take-away" for their students' investment of time and effort to submit an entry may not be sufficiently clear or compelling. The other side of the coin may be evident when one examines criteria for evaluating classroom projects and performances. When a final grade is the only incentive to reach for a higher bar, it's likely that many students will decide that good is "good enough."

Greater agreement between teachers and practitioners on what's worth measuring and the criteria for doing so seems a promising means of engaging more students, more fully—whether in engineering design in particular, STEM disciplines more broadly, or literacy in science and technology and fundamental habits of mind. From the findings presented in this article, the use of the EDPPSR outside of PLTW classrooms (where it is a central component of the senior capstone EDD course) by competition judges represents a positive step in that direction. Members of the higher education community and professionals in various STEM disciplines applied criteria not of their own devising to make judgments about the work of students endeavoring to share their efforts with a wider community beyond their classrooms. The flexible and accessible online process appealed to these volunteers, and the experience fostered respect and increased support for a rising generation young people engaged in STEM.

Given that the vast majority of judges in these competitions had no prior exposure to the EDPPSR—and that all had minimal training—what ought to be striking is not the weaknesses evident in terms of their agreement but rather, the extent to which their scores were within an adjacent range without discussion or moderation. Certainly, from an assessment perspective, the degree of agreement among judges for these three competitions, if representative, would be problematic for any high stakes assessment program. However, as a resource to help make the target for students and their project mentors clear, and to link more detailed criteria to evaluative decisions by judges, the use of the EDPPSR may be regarded as a step in the right direction.

Moving forward, based on judges' survey feedback and evidence distilled from the scores they assigned, a number of observations and recommendations can be drawn:

First, as some of the responses to the judges' survey make clear, the EDPPSR is not without some weaknesses that deserve attention if it is to provide consistently clear and applicable criteria for evaluating elements of the design process. In particular, some of the language in element descriptors was unfamiliar—seemingly to some students, mentors, and judges. Reported weaknesses may be addressed through some fine-tuning of the language of some of the descriptors and/or by providing supplementary resources that include definitions and examples of some of the less familiar concepts and terms.

The best response to the perception among a few judges that the EDPPSR is too cumbersome a scoring tool in relation to the depth of some students' work is probably "it depends." If the sole purpose of a competition is to rate and rank entries, then there are likely to be more efficient ways of conducting the evaluation process. However, in the case of all three competitions, the majority of judges provided notes to support or explain the element scores they assigned. If, as in the case of the KC STEM Alliance, the intent is to provide feedback to students, then the depth and complexity of the EDPPSR may be justified. Understanding the difference between a 2 and a 3, or what prevented judges from assigning a given entry a 5, can very well inform students' future work. As a rule, the more instructionally relevant a score decision has the potential to be, the more detailed the rubric ought to be that is used to make that decision.

Regarding the scope and frequency of score discrepancies, competitions interested in using the EDPPSR or any other instructionally-oriented rubric must ask themselves to what lengths they are willing and able to go to reduce those discrepancies. It is quite reasonable to seriously consider the suggestion made by several judges that training be expanded. At the very least, this might be by creating two modules—one that focuses on using the Innovation Portal to create, access, and score portfolio entries, and another that focuses on the rubric itself, the focused holistic scoring methodology intended (in which score decisions are to be made based on "overall best fit" with one or another of the score point descriptors) and highlights how its structure and language informs score decisions. This protocol might be usefully enhanced by adding a practice scoring module that allowed judges in training (as well as students and teachers, it should be noted) to use the rubric to score a complete portfolio and then compare their own scores with those assigned and explained by a team of experienced raters.

In high-stakes assessment contexts, there is an expectation that raters will perform at a certain level, gauged through qualifying rounds and check sets or embedded validity checks. While this may be overkill for the purposes of assigning scores for a competition, some lessons might be adapted from scoring monitoring in assessment contexts. For example, some guidelines and formulas for dropping outlier scores (high and low) might be implemented. The scores of judges whose scores consistently differed from those of their teammates might be flagged, and overridden at the point when top tier entries were being considered for prizes and commendations.

Since feedback from the competition judges so often veered off topic to discussions of students' lack of understanding of various elements or their apparent disregard for the specifics of element criteria, this topic too deserves more attention. Each of the three competitions made clear to students at the outset that their work would be evaluated based on the EDPPSR, and they were directed to the Innovation Portal links to the rubric and to scored and annotated sample entries. It was clear and unsurprising to judges that those students who independently familiarized themselves with the rubric or were familiar with it from coursework in which it played a role developed stronger portfolios. In order for the EDPPSR to be a meaningful evaluation tool—whether for work produced for the classroom or for a competition—students and their mentors, whether teachers, outside reviewers, or others, need to take the time to become comfortable with, and conversant in, the vision of the design process encapsulated in the EDPPSR. Where it differs somewhat from definitions of design from other contexts, those differences should be discussed and resolved by recognizing the clear parallels. In other words, the labels for various elements and what they address may be different, but the substance is appreciably the same. Students also need to be clear that, as when completing a job or college application or submitting work for publication or commendation, they should familiarize themselves early on and review all available instructions and guidelines. Judges from all three competitions expressed dismay over incomplete projects, missing or incorrectly placed entries, and other weaknesses that could have easily been addressed through more careful attention to competition requirements.

Some of the greatest challenges identified by judges from both the Conrad Spirit of Innovation Challenge and the KC STEM Alliance competitions related to the fact that a rubric originally developed to describe performance in engineering design was being used to evaluate projects which had much, but not necessarily everything, in common with that field. This gave rise to questions, still under consideration by the KC STEM Alliance, about developing a parallel rubric better suited to biomedical research and to concerns expressed by a former Conrad Teacher of the Year about the potential misalignment of the rubric and more research-oriented submissions to the Conrad competition (Halpin, 2016).

For the 2015-16 competition, however, the KC Alliance acknowledged the likely aspects of misalignment and took steps to address them that included directing judges to make small substitutions in wording as appropriate and not including one Element deemed unsuited to biomedical research projects. These decisions may serve to highlight that no rubric—whether the EDPPSR or some other—need be etched in stone. The very nature of evaluative criteria is that they can provide guidelines rather than inflexible rules, and often they can be treated somewhat flexibly and their language and use adapted to fit different contexts. Simplifying the language and structure of the EDPPSR to make it more "student friendly" as a classroom tool, or removing the word "engineering" or changing references to prototype to "experiment" so that design in other disciplines might be considered, are just a few of an infinite number of ways that this particular rubric can be modified, depending upon the audiences and purposes for which its use is intended.

The absence of supplementary resources to support score decisions in other fields than engineering design need not lead to rejection of the rubric either. In fact, successful long-term use of the EDPPSR by a variety of competitions and in a variety of instructional contexts is more likely if stakeholders invest the time to select, annotate, and make available a richer array of sample portfolio entries to illustrate more and less effective portfolio entries in a variety of STEM disciplines and even outside of STEM.

While the focus of this study was the use of a rubric outside the classroom, findings suggest a number of considerations regarding rubric use that apply to instructional contexts as well:

- Consider the fit between the complexity of the rubric being considered and the complexity of the process or product being evaluated, as well as the detail of documentation expected; for example, a complex suite of rubrics like the EDPPSR is far better suited to a semester- or year-long project than one that is the focus of a brief instructional unit.
- Consider modifications to an existing rubric that may be needed for processes and products that do not fit the criteria fully (for example, that address engineering design rather than design more broadly or as defined by disciplines other than engineering).
- Recognize that rubrics exist to be modified or used selectively (that is, utilizing only part of the full suite of rubrics that comprise the EDPPSR) to fit purpose and audience as well.
- Bear in mind that format as well as content may be modified, for example by highlighting or bolding to make key words and terms and the parallelism of criteria across score points more evident.
- Remember that it takes time to master a rubric and applying it to a process or product gets easier after a few attempts. Confidence and competence in applying rubric criteria (whether by students, teachers, judges, or others) may be bolstered by providing not only scored and annotated exemplars but arranging for a guided experience of the scoring process when exposing new users to a given rubric.
- Realize that adaptation of an existing rubric may not be the best way to meet the need for a scoring tool, and it may be more effective to develop a new one. In that case, rubrics that have stood the test of time may serve as useful models.

Although dreams of an advanced placement exam in the form of a standardized portfolio assessment of the design process appear to have faded, aside from its centrality to PLTW's EDD course the EDPPSR has contributed to other instructional and assessment initiatives. It informs Linked Learning units (https://www.cteonline.org/curriculum/outline/engineering-stem-integrated-projects-model/MxP78R) and provided the template for rubrics used by another innovative e-portfolio program for both middle and high school students (Moore et. al., 2016). By no means least importantly, the EDPPSR also survives as a framework for engaging in, and evaluating, the design process through regional, national, and international competitions. Such competitions are dependent, for now, on the Innovation Portal managed by PLTW and require buy-in from a wide community of stakeholders, including but not limited to students, teachers, mentors, and other academics and professionals in the STEM community. It is clear from the experiences of just three such competitions—the KEEN EMpwr pilot, the Conrad Spirit of Innovation Challenge, and the KC STEM Alliance competitions—that refinement to the rubric, to resources provided to all participants, and to the training protocol would do much to increase the long-term viability of these or similar enterprises. The claim that this would be a worthwhile investment of time and resources by academic institutions, non-profit foundations, and industry partners is fueled by the belief that use of a common metric by students and practitioners helps to build bridges from the classroom to the world at large. Such a metric—whether or not it is the EDPPSR—has the capacity to guide students from the realm of academic exercises to more authentic experiences and create pathways that can lead more students to see themselves as members of a wider STEM community.

ACKNOWLEDGEMENTS

This work would not have been possible without encouragement and assistance from the following individuals (listed in chronological order of involvement): Mark Schroll (former Director of Strategic Initiatives, Project Lead the Way), Elliot Mork (Associate Director of Development Initiatives, Project Lead The Way), Thor Misko (Program Director & Team Leader, Entrepreneurial Engineering Program, Kern Family Foundation), Nancy Conrad (Founder and Chairman, Conrad Foundation), John (Jack) Griem (Partner at Carter Ledyard & Milburn LLP), Alyse Cohen (Delaware North), and Ann Zimmerman (Special Projects Consultant and Industry Coordinator for the KC STEM Alliance).

Thanks are also due to the Kern Family Foundation and the National Science Foundation, both of which provided funding for various phases of development and refinement of the EDPPSR. The views expressed in this article are the author's own and do not reflect the view of either organization. Chevron has provided the primary ongoing support for developments and operation of the Innovation Portal.

REFERENCES

Abts, Leigh. 2011. "Analysis of the barriers, constraints and issues for dual credit and/or an Advanced Placement ®pathway for introduction to engineering/design." Paper presented on June 11, 2011, at the annual meeting of the American Society of Engineering Educators. Retrieved July 1, 2013 from www.asee.org/public/conferences/1/author_index/66429

Carr, Mike. 2012. "The Innovation Portal: Gateway to Engineering Design Excellence." Unpublished manuscript available online: https://www.innovationportal.org/uploads/page/2/242519

Conrad, Nancy. Email to Gail Goldberg, March 16, 2017.

Goldberg, Gail. 2014. Revising an Engineering Design Rubric: A Case Study Illustrating Principles and Practices to Ensure Technical Quality of Rubrics. *Practical Assessment, Research & Evaluation*, 19(8). Available online: http://pareonline.net/getvn.asp?v=19&n=8

Groves, James, Rosemary Reshetar, Mark Schroll, and Leigh Abts. 2012. "The development and implementation of a potential AP for engineering design using a rubric based e-portfolio." Conference presentation at 2012 ASEE. Retrieved May 4, 2015 from https://peer.asee.org/25334

Groves, James, Leigh Abts, Gail Goldberg. 2014. "Using an Engineering Design Process Portfolio Scoring Rubric to Structure Online High School Engineering Education." Conference presentation at 2014 ASEE. Retrieved May 4, 2016 from http://www.asee.org/public/conferences/32/papers/10738/view

Halpin, Myra. Phone conversation about experience as a project mentor with Gail Goldberg, January 7, 2016.

Householder, Daniel L. and Christine E. Hailey, Eds. (2012). *Incorporating Engineering Design Challenges into STEM Courses*. National Center For Engineering and Technology Education.

Maryland State Department of Education. 2008. *Maryland's Accountability Assessment Program*, pp. 19-20; retrieved July 11, 2016 from http://www.msde.maryland.gov/NR/rdonlyres/9659F357-134C-4040-AEB9-5972246E764D/19193/AccountabilityAssessmentProgramDec08.pdf

Misko, Thor. Email to Gail Goldberg, December 8, 2015.

Misko, Thor. Phone conversation about the KEEN EMpwr Competition with Gail Goldberg, December 16, 2015.

Moore, Roxanne, Meltem Alemdar, Jeremy A. Lingle, Sunni Haag Newton, Jeffrey H. Rosen, and Marion Usselman. 2016. "The Engineering Design Log: A Digital Design Journal Facilitating Learning and Assessment (RTP)." Proceeding of the 2016 American Society for Engineering Education Annual Conference & Exposition. See https://www.asee.org/public/conferences/64/papers/14726/view

National Science Foundation.2011. Notice of award and abstract retrieved May 4, 2016 from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1118755&HistoricalAwards=false

Ridge, Kirk. Email to Gail Goldberg, March 20, 2017.

Robelen, Erik (2013). "AP Engineering May Be on the Horizon," *Education Week*, March 29, 2013. Retrieved July 1, 2013 from http://blogs.edweek.org/edweek/curriculum/2013/03/ap_engineering_may_be_on_the_horizon.html

Yancey, Kathleen Blake (1999). "Looking Back as We Look Forward: Historicizing Writing Assessment as a Rhetorical Act." *College Composition and Communication*. 50.3: 483-503. Available online at https://faculty.unlv.edu/nagelhout/ENG714f10/YanceyLookingBack.pdf

Zimmerman, Ann. Feedback on manuscript draft, July 22, 2016.

Zimmerman, Ann. "Reflections on 2016 PLTW Senior Showcase." Attachment to May 6, 2016 email to competition judges and others.

AUTHOR

Gail Lynn Goldberg is an educational consultant and researcher with a focus on student assessment. She provides technical support and professional development services to educational institutions, agencies, and organizations in the areas of assessment development across content areas, scoring, and literacy learning. In that capacity, she served as lead consultant on a National Science Foundation grant to explore the validity and reliability of a rubric intended to provide the framework and scoring criteria used to evaluate engineering design portfolios. Dr. Goldberg received her B.A. in 1971 from Queens College and her Ph.D. in English in 1977 from The Graduate Center, City University of New York. First introduced to direct assessment while still a graduate student, her interest in that discipline eventually led her to serve for a decade as an assessment specialist for the Maryland State Department of Education before becoming an independent consultant.

APPENDIX A

Innovation Portal Competitions: Judge's Rubric Review Questionnaire

What is your name? (optional)

What company, institution, or organization do you represent? (optional)

In which Innovation Portal Competition did you most recently participate as a judge and when did you do so?

Were you familiar with the EDPPSR before using it to score entries for this competition? From what contexts (e.g., from a workshop, classroom use, scoring another competition)?

Before actually applying the scoring criteria of the rubric, how were you trained and/or how did you prepare to use the EDPPSR? How long did this take?

How might the process of becoming familiar with the EDPPSR be made easier and/or more effective?

Which (if any) elements did you find more challenging to score (regardless of how easy or difficult it appeared for students to address those elements)? Why?

What did you find particularly helpful when using the EDPPSR to make score decisions?

How might the process of using the EDPPSR to evaluate engineering design process portfolios be made easier and/or more effective?

*Answer ONLY if you previously used the EDPPSR as a competition judge (otherwise, leave blank): Check one choice below to indicate whether it was:

[ ] The same competition in a previous year

[ ] An earlier phase of the same competition

[ ] A different competition

*If a Different Competition was chosen in the question above, Please specify the competition.

*How did this most recent experience using the EDPPSR compare to your earlier experience(s)?

Feel free to provide any additional reactions, observations and/or suggestions: