Abstract

The preliminary steps toward verifying that a critical thinking rubric has meaning and utility within and outside of classroom use have been undertaken. Analyses offer evidence that the rubric may be used to measure the development of critical thinking in the context of chemical engineering design projects from inception of a problem statement to the final project reports. It has been shown to be an especially effective tool for faculty, who take time to develop high interrater reliability, and for student use in rating each other’s projects. In addition, through solicitation of ratings from industry representatives, there are indications that the constructs measured by the rubric have value in a broader, non-academic setting.

Introduction

One key task in developing or evaluating an educational intervention is that of identifying or developing assessment tools that can measure the intended improvements in student learning that will be produced by the intervention. To this end the authors have developed a rubric to measure critical thinking in the specific realm of chemical engineering fluid mechanics and heat transfer design and problem solving. We are certain that the rubric measures things that we, the educators who developed it, value for assigning grades and assessing competence. However, from the perspective that the purpose of undergraduate engineering education is to prepare students for their future careers, it is important to determine whether the skills and thinking characterized by the rubric are similarly of interest to our student’s potential future industrial employers.

Specific need for validating our rubric

Our rubric is based on a critical thinking rubric (CTR) developed by an assessment group at Washington State University (WSU). The original rubric is broad and multi-disciplinary, has been used throughout WSU (Brown, 2004), and has been a major assessment tool for an NSF grant at another institution (Damron & High, 2008). This original rubric is based on formal analysis, which is rooted in classical rhetoric, and has concurrent validity with multiple definitions of critical thinking (Facione, 1990, 2000; Paul & Elder, 2005; Toulmin, 1979). Though we have modified the CTR in a way that maintains a logical connection to the construct validity of the original and performed the modification in a way that we believe builds content and expert validity for the purpose of assessing critical thinking in our particular course, we are only meeting a portion of the broader need. We have not ensured that the measurement matters to the future employers of the students. Specifically this is the sub-category of content validity known as content relevance (Chatterji, 2003); some frameworks also refer to this as criterion related validity (Moskal & Leydens, 2000).

In addition to relevance, we need to examine issues that relate to the usefulness of the rubric, both as a teaching tool (Wiggins, 1998; Wiggins & McTighe, 2005) and as a tool that can be distributed and used beyond the pool of individuals who helped with the development. Are there systematic biases in between different groups of raters? Is the measurement credible to all stakeholders? Is it both honest and fair?

In order to begin addressing these questions, the authors undertook a study to utilize our CTR as a major assessment component in a project based junior level chemical engineering course on fluid mechanics and heat transfer. The rubric was used for instructor assessments of student work at various milestones during the semester. As a means to help them understand the use and meaning of the rubric, students also used the rubric to assess a student project from the previous semester. The final project reports were rated by the instructors for grades, the students as their final, and by a group of alumni who were willing to take part.

In this article we begin to examine the content relevance of our rubric by examining the differences and similarities between how the different groups rated the final project reports. These results, and supporting survey responses, provide some insight into probable biases between groups, the credibility of the rubric to different groups, and the portability of the rubric to new users.

Background

Why Critical Thinking?

For at least the past 20 years academics have been claiming that engineers need more training in what are frequently referred to as “soft skills” (Adams et al., 2011; Dickson & Grant, 2007; Felder, Brent, & Prince, 2011; Litzinger, Lattuca, Hadgraft, & Newstetter, 2011; Varma, 2003). Engineering educators generally think of soft skills as the non-technical items in ABET criterion 3 (ABET, 2003), which include items such as “an understanding of professional and ethical responsibility” and “the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental and societal context.” These claims have been echoed by industry, as evidenced by a recent study conducted by the American Association of Colleges and Universities (Peter D. Hart Research Associates, 2008), surveying 301 employers, which shows a frustration with the “broader skills” of recent graduates. Boeing has taken this frustration a step farther and implemented a process that aligns Boeing’s hiring practices with their internal evaluation data to track back to college programs that provide them with successful hires (Baskin, 2008). Critical thinking is a key component that underlies many of the soft skills with which we are concerned (Anderson, Howe, Soden, Halliday, & Low, 2001) as it is the base upon which other soft skills such as teamwork, leadership, communication, engineering solution context, and professional and ethical considerations must be built. Without critical thinking all other skills, though they may be present to an extent, will lack the solid thoughtful and rigorous foundation upon which judgements are based.

Critical thinking is a broad term and multiple definitions have been proposed. For example, Facione (1990) gives a definition for critical thinking as follows:

“a purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based . . .”

Looking at this definition of critical thinking, we can see that it also describes, in broad terms, the process by which engineers solve ill-structured problems; namely coming to a solution (judgment) by a process of interpretation, analysis, evaluation and inference using principles (concepts), standards (methods), regulations (context),
and client criteria. Ill-structured problems are problems that do not have a single correct answer. In engineering this includes design and analysis of an existing design or proposal. For this reason problem solving has been selected as an appropriate, domain specific sub-set of critical thinking.

Why Rubrics?

Though multiple standardized tests exist that measure aspects of critical thinking (Anderson et al., 2001; Facione, 2000; King, Wood, & Mines, 1990; Terenzini, Springer, Pascarella, & Nora, 1995), what we are concerned with here is measuring critical thinking as shown in a student’s written products. In education assessment terms this is a “performance” and rubrics are especially well suited for the systematic assessment of performances relative to criteria (Wiggins & McTighe, 2005). This allows us to embed the assessment in a task that resembles something that might be done by a professional in the field (Moskal & Leydens, 2000).

From a formative assessment perspective, rubrics allow for feedback that is focused and directly related to the assignment. This provides an opportunity to explicitly show a student where and why they went wrong, which provides an opportunity for them to correct their understanding. It is well established that formative assessment has many benefits (Bransford et al., 2000). Having a formal tool that provides a framework and structure for feedback should be helpful for both students and teachers.

Rubric History and Development

As was previously mentioned, our rubric is a modification of a broader, existing CTR. The original rubric measures critical thinking in 7 dimensions on a 3-point scale; emerging, developing and mastering. The dimensions are: identifies and summarizes the problem/question; identifies the student’s own perspective, hypothesis, or position; identifies and considers other salient perspectives and positions; identify and assess key assumptions; identifies and assesses the quality of supporting data/evidence; identifies and considers the influence of context; identifies and assesses conclusions, implications, and consequences. With a well-designed rubric already existing, it may appear that there is no reason to make modifications; however, the authors believed that a more specific rubric, which used language and examples from the discipline of chemical engineering, would be more accessible to and usable by chemical engineering students and faculty.

Details on what was changed and why may be found elsewhere (Golter, Abdulla, Thiessen, Brown, & Van Wie, 2010; Golter et al., 2008; Golter, Van Wie, & Brown, 2007; Thiessen et al., 2009), however a brief summary is presented here. The revisions were done over a series of 4 years, and were undertaken by a group consisting of 3 chemical engineering professors and 1 chemical engineering graduate student, with guidance from an educational assessor. The instructional subject matter expertise of the group would ensure that the modified rubric has utility for measuring constructs that matter for academic use in engineering education. Furthermore, this revision was done in a way that maintained a logical connection with the parent instrument and the research pool it drew from. For example, the categories and criteria of the modified rubric maintained alignment with Facione’s broad definition of critical thinking.

The modified rubric, appearing in Appendix 1, contains 9 dimensions: problem identification, fluid mechanics principles and equations, heat transfer principles and equations, fluid mechanics assumptions, heat transfer assumptions, other assumptions, equipment specification, solution quality, and organization/communication. These 9 dimensions correspond to a subset of the dimensions of the original rubric. In both cases, the first dimension is problem identification. The second dimension of the original rubric, presenting the student’s own perspective became more specific to engineering problems in the form of presenting students’ solution methodology, which became both “Principles and equations” and “Equipment specification” dimensions in the present rubric. This particular dimension, along with assumptions, was made more analytic by separating it out by subject. “Identify and assess conclusions, implications and consequences” became “Solution quality.” The dimension of “Identify and consider other salient perspectives and positions” was incorporated into the “Equipment specification” section through metric descriptors “Thorough evaluation of alternatives” and “Solutions . . . can be extended to other situations”, the “Identify and assess the quality of supporting data/evidence” dimension was distributed among the dimensions that come out of presenting the student’s perspective, especially in assessing assumptions, solution quality and conclusions, and the “Identify and consider the influence of context” dimension was incorporated into the new rubric’s “Solution quality” dimension as represented in the rubric descriptor stating “Solutions are accurate, appropriate, thorough, and clearly linked to design problem.” “Organization and communication” was added as a rubric dimension in order to encourage raters to separate the student’s writing from their thinking.

The modified rubric works on a scale from 0 to 6 with half, 0.5, point increments rather than using the original rubric’s 3 point scale. Conceptually this adds “absent” to the “emerging / developing / mastering” scale and further subdivisions for finer differentiation in rating student assignments. We further define a 4 as the desired competence level of graduating undergraduate engineers which if achieved places them in “developing”, but on the cusp of “mastering” range. This allows alignment with the Educational Testing Services (ETS, 2010) convention of using an absolute scale, and anchoring at 2 points, absent (0) and competent (4), the understanding we would expect of the average graduating senior. The scale ranges from 0 to 6, which represents a range from little to no understanding at levels 0 to 1 up to complete mastery, level 6. The choice of level 4 as an anchor is an important one as it calibrates the rubric at a level closest to the learning expectations in the undergraduate curriculum, and gives room for raters to assign a score for those who go over and above the average, somewhere above average performance but presumably less than the level of an experienced engineer.

Methods

Overview of Class and Assignments

The rubric was used in a required junior level chemical engineering course. This course, Fluid Mechanics and Heat Transfer, is two credits, is the second course in our “transport series”, and follows a course titled Transport Phenomena. At this point the students have had the relevant theory in both momentum and energy transport. The course in question is focused on the application of these theories, specifically topics such as sizing pumps and heat exchangers. The instructor has, for many years, included a small group design project as a significant portion of the course. The class was split into groups using fairly standard criteria of mixing ability level, as measured by entering GPA, avoiding creating groups with solitary female or minority students, and trying to ensure that group members had compatible schedules. There were seven groups of five students and one group of four, for a total of 39 students.

For this semester, we used a unique semester-long team project-based pedagogy where we promoted learning based on goals, procedures and outcomes stated in the syllabus, in-class interactive exercises to inform project efforts, the CTR and its use, and well written, but not perfect, sample project artifacts from a previous course offering. We emphasized from the first day of class the program level outcome of creating chemical engineers that individually understand the conceptual and practical aspects of the discipline with the professional skills needed to develop a rigorous multi-component process design in a team setting. Students were told they would be treated as practicing engineers in industry who by the end of the course should be competent to analyze and design fluid mechanics and heat transfer systems in a simulated “real world” environment.

An ill-structured project statement was given in the syllabus to “design a fluid mechanics and heat transfer system to solve a practical problem: e.g., a hot waste stream from an ion exchange regeneration process needs to be treated before being discharged.” However, a specific set of learning outcomes, Insert 1, was also given in the syllabus and students told they were to meet all of them. To buttress team efforts on a project proposal due at the end of the first week an example problem statement from a successful past project (see C2J2T team project proposal in Appendix 2) was discussed in class through an instructor-facilitated team-based think-pair-share exercise before the due date. The CTR was also given in the syllabus as a basis on which all project aspects would
be rated. After submission of the first project proposal draft a half-page set of strengths and weaknesses were given to each team to guide improvement on an improved project proposal due two weeks later. Assignments continued in this vein with a recursive continuous improvement strategy.

The next set of assignments focused on using the CTR in rating an entire prior project report (in this case a design of Willy Wonka’s chocolate waterfall), in guiding team projects, in rating each other’s projects, and in instructor rating and feedback on assignments. To promote adherence to the CTR-centered approach, a graduated scale was given in the syllabus with progressively higher ratings needed on each assignment as the semester progressed. General statements were made on the criterion for each letter grade, e.g., an “A” would be earned when “given an open-ended problem, students demonstrate the ability to define the problem, determine the underlying principles that apply, make reasonable assumptions, come up with a solution and then discuss the implications of that solution.” To gain familiarity with the CTR and understand project expectations, in-class scaffolding was provided by the instructor through another guided think-pair-share on aspects of the “Willy Wonka” project, following which individual students performed and turned in their own rigorous analyses and CTR ratings. Feedback was given on student ratings, again using the CTR to rate how well they did in rating the case study project. A mid-semester team project was submitted and rated by the instructor and TAs based on the CTR. Specific feedback including both strengths and improvements was provided on each metric in the CTR for improving the next and final project submission. In lieu of a final exam, each team was given the task of rating another team’s final project report, again using the CTR. To differentiate effort among team members, and motivate individuals to participate in the team effort, each student rated all team members, including themselves, on their level of participation in the rating process—this served to provide a multiplier to the team score to provide a method for assessing contributions resulting in a differentiated grade for each student on a team.

Credit for the various assignments was given as follows. The project was worth a total of 60% and was a group assignment with a graduated increasing percentage weighting correlated to the difficulty of the assignment and expected level of team development within the semester. Project Proposal I was worth 2%, Proposal II 8%, initial project report 10%, and final project report and presentation 40%. The first “case study” or rating of a prior project was worth 10% and was an individual assignment. Standard individual homework assignments made up 20%, and the final “case study,” where the student groups evaluated another group’s design made up the final 10% of the course grade.

ETS Convention in Rubric Ratings

We follow the convention laid out by Educational Testing Services (ETS, 2010) when using the rubric. In this convention typically each paper is rated by at least two people. Ideally raters would sit down as a group to rate all of the papers, with everyone rating one paper together at the beginning and at 1-hour intervals until the end. The purpose of being to establish a norm, or common understanding of the rubric. Periodic checks throughout the rating session provide an indication of how well the norm is being maintained. These norming sessions involve everyone rating a paper individually, followed by a discussion intended to reconcile differences in the ratings. In the standard ETS convention a difference is any rating in which the raters have assigned an average score that is more than 1 point from another rater’s, or which crosses the line between developing and proficient. This was done in lieu of a final exam and had a significant impact on the student course grades. It is possible that this influenced students to be extra critical of each other’s work. Also, since each group rated only one other group’s paper, between group IRR statistics are not possible.

Ratings by students

The second case study provides a data point for students using the rubric to rate the same set of assignments the instructors were rating. At this point, the very end of the semester, the students had been working with the rubric throughout the semester and were familiar with it. This was done in lieu of a final exam and had a significant impact on the student course grades. It is possible that this influenced students to be extra critical of each other’s work. Also, since each group rated only one other group’s paper, between group IRR statistics are not possible.

Ratings by employers

With the assistance of the college development office, we recruited a group of eight alumni to rate the students’ final design presentation and to provide feedback, using the questions shown in Table 1, on the applicability of both the assignment and the rubric to an industrial career. These alumni ranged in experience from 20 to 50 years and represented industries including chemicals, petroleum, pharmaceuticals, environmental cleanup, aluminum and pulp and paper. Four of the alumni had last held executive level positions, and the remaining four held management positions. These alumni represent a highly experienced group who should be accustomed to evaluating designs and proposals. Their insights here should provide a comparison between industrial and academic expectations for this type of assignment.

The papers were distributed so that each paper was rated by at least two alumni. This, however, required two rounds of recruiting due to alumni either dropping out of the study or not responding. In spite of the extra recruiting, we were only able to collect one numerical rating for one of the papers, which were somewhat corroborated by the written comments of a second rater. Since this was a geographically dispersed group, we made no attempt to arrange a norming session. The ratings were collected using an online survey system.
### Statistical Methods – Reliability

The previous discussion of what constitutes a "difference” leads directly to one of the key statistics for this type of rubric use, Inter-Rater Reliability (IRR). Quite simply, this is the ratio of the number of comparisons of the average score that agree within 1 divided by the total possible number of comparisons. For a given paper, each rater can compare scores with each other rater. In the case of four raters, this gives six possible comparisons. Mathematically the total number of possible comparisons is as follows:

\[ \text{Number of comparisons} = \sum_{i=1}^{4} \text{Number of Raters} i - 1. \]

By convention, an IRR of 70% or greater is satisfactory. The IRR can be calculated for ratings both within group and between groups. By grouping the ratings we can see whether we had reliability between the different groups of raters. For example, if we group the alumni and academic ratings the resultant IRR includes the reliability between these two groups. IRRs within and across groups highlight differences in opinion on the relative strength of an argument being made in a student paper. Where IRRs are in agreement it may imply strong agreement on meaning of the rubric criteria or at least that students have met the criteria, albeit that different raters may have seen the criteria met by differing statements in a report. There is a possible source of bias in this due to the larger sample size for the academic raters. It is likely that this reduced the influence of other groups’ ratings when calculating a between group IRR.

### Qualitative Methods – Credible to all stakeholders

To examine whether our rubric is part of a credible assessment, we need to determine if what we are measuring matters and makes sense to the stakeholders in our students’ education. In this regard our study was exploratory. Along with rating a small number of papers, alumni were given a brief survey requesting their feedback on both the rubric and the project assignment. The purpose of this survey was to assess the content relevance of the assignment and rubric to an industrial career. This was appended to the survey page for collecting the alumni’s ratings. This survey was not rigorously developed or administered, and is meant to provide direction for further inquiry rather than providing definitive answers. Table 1 shows the survey questions and response choices.

As another exploratory companion to the alumni survey, we held a brief informal discussion with a trio of students who were involved with another aspect of this research project. During this discussion we probed their perceptions of the rubric and the assignment.

### Results

#### Overview

The first, and easiest to assess, set of data to examine is the scores from the second project submission. This may be seen graphically in Figure 1. There is a visually apparent trend toward the students rating more harshly and alumni rating more easily compared to the instructors. However, as can be seen from the error bars, the differences are not statistically significant, possibly due to the small sample size (n=8). At this small sample size, a difference would need to be quite large in order to be significant. As none of the differences between any of these groups even approach one standard deviation, no further statistical testing was required to determine that the differences

<table>
<thead>
<tr>
<th>Table 1: Survey given to the alumni raters regarding the rubric and assignment. These results were given a numerical value ranging from 1 to 6 in intervals of 0.5.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table 1:</strong> Survey given to the alumni raters regarding the rubric and assignment. These results were given a numerical value ranging from 1 to 6 in intervals of 0.5.</td>
</tr>
</tbody>
</table>

### Figure 1: Average ratings by each group of raters, sorted from highest to lowest, with 1 standard deviation error bars.
could be explained by chance. When phrased as grades the differences are more striking with alumni giving five scores of A- or better (ranging from A+ to C, averaging A-) academics giving only two As (ranging from A to D+, averaging B), and students giving only one A- (ranging from A- to D-, averaging B-).

Reliability Results

The IRRs within each group may be found in Table 2 which, as shown, range from 86%-100% for reconciled values with improvement from the 77% for initial use with the "Case Study" to 100% for initial ratings of the Project 2 and final "Case Study". An exception is the alumni IRR of 36%. Where the data are available the initial IRR is included. All of the IRRs were calculated using the ETS IRR of 36%. Where the data are available the initial IRR will likely not appreciably change over time. However, if a rater is properly trained in the rubric’s use, their ratings may be important for raters to re-norm if significant time has elapsed before another assignment is rated. However, over the course of a semester we were able to maintain this higher, within 0.5, level of accuracy.

Credibility of Stakeholders Results

Alumni answers to the survey questions are summarized in Table 5. In addition to the numerical summary, average and range, there is a pair of extra columns that reflect a misunderstanding with regard to the survey. Though the survey was specific to the rubric rather than the paper being rated, in several instances, an individual rater gave different scores for questions that should have had consistent answers regardless of the student paper being rated, since the questions were focused on the rubric. The consistency column is added to show the percentage of times where the raters gave the same score across different papers, and the abstention reflects that some questions were randomly skipped.

When asked to give their reasoning some of the alumni responded with comments that were clearly directed at the reports.

"Report is well organized and clearly discussed. Engineers usually have trouble writing and explaining with the written word.”

Others indicated that the rubric was a good starting point for this type of assignment.

"The rubric elements plus economic analysis would also be required.”

"The rubric covers the essentials in the assessment of a chemical engineer’s design and does an excellent job of assessing their needed capabilities to start. . . . recommend a specific addition to the rubric covering the specific writing skills and not hidden in the communication assessment.”

Finally we asked what they would do to improve the assignment. Responses largely dealt with issues that are, currently at least, outside of the scope of our course.

"Addition of economic analysis and formal report writing. These are essential for opportunity for advancement.”

"Pictures and drawings help. A course in PFD [process flow diagram] software would help their ability to communicate intent and technical thinking. This could be a simple part of the orientation course.”

Two of the suggestions would be implementable in our course. First adding an abstract or executive summary to the report:

"Organization of their thinking … start by adding an abstract … “

This would also address many of the comments that came up earlier in the ratings of the each report, and help support the organization and communication dimension of the rubric. Lack of an executive summary or detailed flow...
had the opportunity to work together multiple times over

perceived as being interested in and willing to help the de-

which consisted of senior level to retired people who were

group to recruit participants, we had a group

ratings and students the easiest, the data in figure 1 shows

One student mostly dominated the discussion. The manner in

“... many times work in industry involves investigation of pro-

reinforce this. The students stated that they did not want to

Unsurprisingly, the results in Table 2 show that we

We had expected to see the alumni give the harshest

We had attempted to meet this goal by having the students

evaluate a past report at the onset of the course, as well as each other’s reports at the conclusion of the course. This por-

discussion about the intended audience for their reports

The results in Table 4 do show the 100% IRR on the

“... was our first ChemE project!”

One student claimed that using the rubric also left the students

assignment and rubric help build skills necessary for an in-

This feedback also lends support to the indications that the

Credibility Discussion

Based solely on the average responses in Table 5, it

Further Implications

Since the work presented in this paper is an initial effort to

validate a critical think rubric in the chemical engineering context, we have demonstrated success in terms of inclusion of specific categories for application in fluid mechanics and heat transfer, achieved an outstanding level of agreement between trained faculty raters to the extent that in imple-

mentations of the rubric at the final report stages there was

The results are roughly indicative that we have succeeded

in examining a construct that is important to an industrial career. Based solely on the average responses in Table 5, it

appears that the rubric measures a set of skills which are

important to an industrial career and that the assignment

provides a good opportunity for students to practice these skills. However, taking the abstentions and the consistency into account, we are not justified in making a strong claim. Based on the range of scores though, we can say that indi-

cations are that the rubric probably measures a set of im-

portant skills and the assignment seems to provide a good opportunity to practice the measured skills. The written comments, especially those regarding the rubric as a good

Starting point for design, support this claim. It is clear that in

follow up work, questions regarding the quality of the rubric and assignment need to be given separately after alumni have had a chance to rate some papers.

In some ways it is more difficult to assess the broader meaning behind the question of “at what level would you hire...” The average result of 3.1 is on the cusp between internship and entry level, 3.0 being the highest score for internship and 3.5 being the lowest score for entry level. One would expect students to easily be ready for an in-

tership by the end of their junior year, in this respect the

students are meeting our expectations and in some cases exceeding them, specifically four of the eight groups rated at

entry level. Based on this, the assignment appears to pro-

vide students a good opportunity to showcase their skills. This feedback also lends support to the indications that the assignment and rubric help build skills necessary for an in-

Disscussion

Overview

We had expected to see the alumni give the harshest

differences. However, all of these IRRs are above the

minimum criteria of 70%. As an interesting aside, all but one

of the differences noted had a discrepancy between compe-
tent (4.0) and developing (3.5), while within 0.5 cross the

certainty anchor of 4.0. The one remaining difference

was a difference of 1, and would not have counted if not

for the student-requested change from ETS convention to a

tighter 0.5 scale for better differentiation of grades. The IRR

breakdown reveals the rubric may be used effectively as is,

though the ideal for what students should graduate capable

doing is less well developed when it comes to “Other As-

sumptions” such as safety and environmental concerns. This

may be taken into account in future refinements of the rub-

One student also asked if the alumni had any oppo-

nents as well.

between industrial members and perhaps with faculty

members as well.
Conclusions

The rubric we have developed shows indications of being a valid tool for measuring a student’s critical thinking with regard to engineering design and problem solving in a manner that is consistent with the needs of industrial employers. Unfortunately, this statement must be tempered with the knowledge that our efforts to validate this instrument in the broader context of industrial needs is only beginning and we have only established a preliminary validity in this regard. Aside from this exploratory conclusion, the data for the academic raters clearly show that providing the raters an opportunity to get together and discuss the rubric and their ratings very significantly enhances reliability in rubric measurements, as starkly evidenced by the 100% IRR among academic raters compared to the 36% IRR among alumni.

A brief analysis of combined reliabilities indicates that academics probably occupy an in-between perspective between that of students and alumni. When combined with the student comments, which indicate an inaccurate view of the workplace, this indicates an opportunity for faculty to address student misconceptions regarding what the workplace will be like. Based on the reliability results however, it is also possible that student and academic perspectives may be closer together than academic and alumni perspectives. These results are ambiguous due to differences in sample size and the very high reliability among academic raters potentially masking the results.

It is clear that much more needs to be done. To begin with, we need to sample a broader cross-section of alumni. Selecting two papers from the set used in this study and soliciting ratings from a large group of alumni, ranging from five to 50 years in experience, would determine whether there are effects based on experience level. The survey requesting feedback specifically on the rubric and assignment needs to be given separate from the rating collection. Both this survey and either a formal interview or survey with students need to be formally developed in order to draw stronger conclusions from the data.

Acknowledgements

The authors acknowledge support from the US National Science Foundation CCLI Grants DUE-0618872, 1023121, 143267, the Norcliffe Foundation, and a World Bank STEPB Grant through collaboration with the chemical engineering Department at Ahmadu Bello University, Nigeria, all on advancing and assessing use of new pedagogies fostered by use of Desktop Learning Modules. The authors appreciate valuable comments and guidance from members of the Washington State University Engineering Education Research Center, namely Profs. Shane Brown, Denny Davis, and Richard Zollars, and Dr. Ashley Ater-Kranov. We further acknowledge assistance from TAs Baba Abdul and Ashfaq Anshy as well as from WSU Engineering Development Office Director Don Shearer.

References


ETS. (2010). Retrieved from http://www.ets.org portal/site/ets/menueitem.3a88fe2842a2ada76c5e5a10c3921509?vgnextoid=85b65784623f4010VgnVCM100000295590RCRD


Paul Golter did his B.S. at the University of Idaho, his M.S. and Ph.D. along with postdoctoral research at Washington State University, and held an adjunct lecturer position at the University of Idaho. His research has been focused on developing and assessing “Desktop Learning Modules”. He is currently a mechanical engineering lecturer at Ohio University.

Gary Brown is a Senior Fellow with the Association of American Colleges and Universities (AAC&U) and Academic Director of the international ePortfolio organization, AAEEBL. With AAC&U, Dr. Brown recently served as the assessment lead exploring the potential of Lumina’s Degree Qualifications Profile (DQP) as a way to effectively mediate transfer. The initiative engaged 13 states including two and four-year institutions. At AAEEBL, Dr. Brown co-directs the research initiatives and co-chairs the Vendor Advisory Council. Dr. Brown serves on advisory boards for Quality Matters and Campus Technology. His teams have received six best research awards and two best publication awards.

Bernard Van Wie did his B.S., M.S. and Ph.D. along with postdoctoral research and visiting lecturer positions at the University of Oklahoma. He has been on the Washington State University (WSU) Voiland School of Chemical Engineering and Bioengineering faculty for 33 years, 19 years of which have been focused on developing and assessing “Desktop Learning Modules”. He won WSU’s 2009 Marian Smith and 2016 inaugural Teaching Innovation Awards. His technical research is in the areas of biomedical engineering and biomass conversion.
<table>
<thead>
<tr>
<th>Dimension</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Problem / Question</td>
<td>Does not identify a specific question or necessary and appropriate system or systems. The question or system, if identified, is confused or simplistic.</td>
<td>Identifies a somewhat focused question that is interesting but not particularly challenging or is simplistic, tends to ignore essential constraints.</td>
<td>Identifies a focused, unique, original question that is challenging and well-defined.</td>
<td>Totally understands constraints.</td>
<td>Thoroughly understands constraints.</td>
<td>The question is thoroughly and characterized.</td>
<td>Score</td>
<td></td>
</tr>
<tr>
<td>2. Fluid Mechanics Principles / Equations How chosen/appropriateness</td>
<td>Analysis of fluid mechanics principles is absent, incorrect or inappropriate relative to the presenting problem—or insufficiently related to the challenge the project entails.</td>
<td>Analysis of fluid mechanics principles is essentially correct; perhaps some is off target or barely related.</td>
<td>Analysis of fluid mechanics principles is essentially correct; perhaps some is off target or barely related.</td>
<td>Analysis of fluid mechanics principles is thoroughly and correct.</td>
<td>Analysis of fluid mechanics principles is thoroughly and correct.</td>
<td>Interpretation is well integrated with other chemical engineering principles, sources and professional perspectives.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth of Use / Integration Comprehensiveness</td>
<td>There is little or no synthesis from fluid mechanics principles, or there may simply be a restatement of inherent facts.</td>
<td>Interpretation is adequate and clear, though perhaps not fully integrated with other sources or perspectives. Barely extends, if at all, beyond rudimentary exploration.</td>
<td>Interpretation is adequate and clear, though perhaps not fully integrated with other sources or perspectives. Barely extends, if at all, beyond rudimentary exploration.</td>
<td>Interpretation is well integrated with other chemical engineering principles, sources and professional perspectives.</td>
<td>Interpretation is well integrated with other chemical engineering principles, sources and professional perspectives.</td>
<td>Interpretation is well integrated with other chemical engineering principles, sources and professional perspectives.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Heat Transfer Principles / Equations How chosen/appropriateness</td>
<td>Analysis of heat transfer principles is absent, incorrect or inappropriate relative to the presenting problem—or insufficiently related to the challenge the project entails.</td>
<td>Analysis of heat transfer principles is essentially correct; perhaps some is off target or barely related.</td>
<td>Analysis of heat transfer principles is essentially correct; perhaps some is off target or barely related.</td>
<td>Analysis of heat transfer principles is thoroughly and correct.</td>
<td>Analysis of heat transfer principles is thoroughly and correct.</td>
<td>Interpretation is well integrated with other chemical engineering principles, sources and professional perspectives.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth of Use / Integration Comprehensiveness</td>
<td>There is little or no synthesis from heat transfer principles, or there may simply be a restatement of inherent facts.</td>
<td>Interpretation is adequate and clear, though perhaps not fully integrated with other sources or perspectives. Barely extends, if at all, beyond rudimentary exploration.</td>
<td>Interpretation is adequate and clear, though perhaps not fully integrated with other sources or perspectives. Barely extends, if at all, beyond rudimentary exploration.</td>
<td>Interpretation is well integrated with other chemical engineering principles, sources and professional perspectives.</td>
<td>Interpretation is well integrated with other chemical engineering principles, sources and professional perspectives.</td>
<td>Interpretation is well integrated with other chemical engineering principles, sources and professional perspectives.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Fluid Mechanics Assumptions</td>
<td>There is little or no synthesis from literature, or what is presented is incoherent or patchwork together without explanation or demonstration of underlying fluid mechanics assumptions. Assumptions tend to be confused or perhaps contradictory. Key aspects of the challenge are neglected.</td>
<td>Application of material discerned from literature or outside reading is adequate in scope and accuracy, though perhaps slightly confused at times or partially inaccurate. Sometimes questions and attempts to support the validity of assumptions.</td>
<td>Application of material discerned from literature or outside reading is adequate in scope and accuracy, though perhaps slightly confused at times or partially inaccurate. Sometimes questions and attempts to support the validity of assumptions.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Heat Transfer Assumptions</td>
<td>There is little or no synthesis from literature, or what is presented is incoherent or patchwork together without explanation or demonstration of underlying heat transfer assumptions. Assumptions tend to be confused or perhaps contradictory. Key aspects of the challenge are neglected.</td>
<td>Application of material discerned from literature or outside reading is adequate in scope and accuracy, though perhaps slightly confused at times or partially inaccurate. Sometimes questions and attempts to support the validity of assumptions.</td>
<td>Application of material discerned from literature or outside reading is adequate in scope and accuracy, though perhaps slightly confused at times or partially inaccurate. Sometimes questions and attempts to support the validity of assumptions.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Other Assumptions</td>
<td>There is little or no synthesis from literature, or what is presented is incoherent or patchwork together without explanation or demonstration of underlying heat transfer assumptions. Assumptions tend to be confused or perhaps contradictory. Key aspects of the challenge are neglected.</td>
<td>Application of material discerned from literature or outside reading is adequate in scope and accuracy, though perhaps slightly confused at times or partially inaccurate. Sometimes questions and attempts to support the validity of assumptions.</td>
<td>Application of material discerned from literature or outside reading is adequate in scope and accuracy, though perhaps slightly confused at times or partially inaccurate. Sometimes questions and attempts to support the validity of assumptions.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Equipment Specification</td>
<td>There is little or no synthesis from literature, or what is presented is incoherent or patchwork together without explanation or demonstration of underlying heat transfer assumptions. Assumptions tend to be confused or perhaps contradictory. Key aspects of the challenge are neglected.</td>
<td>Application of material discerned from literature or outside reading is adequate in scope and accuracy, though perhaps slightly confused at times or partially inaccurate. Sometimes questions and attempts to support the validity of assumptions.</td>
<td>Application of material discerned from literature or outside reading is adequate in scope and accuracy, though perhaps slightly confused at times or partially inaccurate. Sometimes questions and attempts to support the validity of assumptions.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td>Application of material discerned from literature as well as from outside reading is complete and accurate.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Solution Quality</td>
<td>Solutions are missing or inaccurate/irrelevance. The implications of the solutions are absent.</td>
<td>Solutions are reasonable, though perhaps incomplete or limited. Solutions relate to the design problem and arise from the analysis presented, though there may be gaps or inconsistencies. May include speculation about implications -- mostly plausible, but not necessarily reasonable and useful.</td>
<td>Solutions are reasonable, though perhaps incomplete or limited. Solutions relate to the design problem and arise from the analysis presented, though there may be gaps or inconsistencies. May include speculation about implications -- mostly plausible, but not necessarily reasonable and useful.</td>
<td>Solutions are accurate, appropriate, thorough, and clearly linked to design problem. Solutions and recommendations are qualified, balanced and can be extended to other situations.</td>
<td>Solutions are accurate, appropriate, thorough, and clearly linked to design problem. Solutions and recommendations are qualified, balanced and can be extended to other situations.</td>
<td>Solutions are accurate, appropriate, thorough, and clearly linked to design problem. Solutions and recommendations are qualified, balanced and can be extended to other situations.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Organization and Communication</td>
<td>Presentation of problem, analysis, solution and interpretation seems haphazard, inconsistent, or misleading; one or more elements may be missing or confused. Organization of ideas / multiple errors obscure meaning, and may mislead or misdirect audience. Many parts seem difficult for the audience to follow. Communication style, written and/or oral, is not appropriate to this discipline, or is confusing. Does not use language of the discipline, or uses it incorrectly. Frequent errors may obscure ideas.</td>
<td>A reasonable progression from problem to analysis, solution and interpretation, linked to the problem and solution.</td>
<td>Presentation is appropriate for intended audience, though there may be occasional gaps, errors, or inconsistencies which require effort from audience in order to understand. Communications style, both written and oral, is appropriate to this discipline, though not at a professional level. Communication is adequate. Incorporates some language of the discipline, though incompletely. Some errors may distract audience.</td>
<td>Progression from problem to analysis, solution and interpretation is concise, creative, and clearly links problem to solution and implications. Needs and interests of intended audience effectively inform presentation’s approach and organization. Audience seems well able to follow the presentation. Communication style, both written and oral, is appropriate to the discipline, and is polished, professional, and virtually error free. Uses language of the discipline fluidly and effectively.</td>
<td>Progression from problem to analysis, solution and interpretation is concise, creative, and clearly links problem to solution and implications. Needs and interests of intended audience effectively inform presentation’s approach and organization. Audience seems well able to follow the presentation. Communication style, both written and oral, is appropriate to the discipline, and is polished, professional, and virtually error free. Uses language of the discipline fluidly and effectively.</td>
<td>Progression from problem to analysis, solution and interpretation is concise, creative, and clearly links problem to solution and implications. Needs and interests of intended audience effectively inform presentation’s approach and organization. Audience seems well able to follow the presentation. Communication style, both written and oral, is appropriate to the discipline, and is polished, professional, and virtually error free. Uses language of the discipline fluidly and effectively.</td>
<td>Progression from problem to analysis, solution and interpretation is concise, creative, and clearly links problem to solution and implications. Needs and interests of intended audience effectively inform presentation’s approach and organization. Audience seems well able to follow the presentation. Communication style, both written and oral, is appropriate to the discipline, and is polished, professional, and virtually error free. Uses language of the discipline fluidly and effectively.</td>
<td>Progression from problem to analysis, solution and interpretation is concise, creative, and clearly links problem to solution and implications. Needs and interests of intended audience effectively inform presentation’s approach and organization. Audience seems well able to follow the presentation. Communication style, both written and oral, is appropriate to the discipline, and is polished, professional, and virtually error free. Uses language of the discipline fluidly and effectively.</td>
</tr>
</tbody>
</table>
Due to the nature of Christmas, elves find themselves incredibly busy throughout the working year. Because of this, the plants should be running at optimum efficiency in all aspects. Presently, the production of tea and cocoa has a few inefficient aspects. The beverages are prepared in a separate building from the factory creating a gap in work whenever the elves require something to drink. It also requires them to trek through the polar winds to acquire a hot drink, a counterproductive process.

To fix this, we propose a means of transporting the drinks directly from the place of production to the factory itself. We will use insulated tubing located above the ground. The process will begin with heating water to roughly 136 degrees Fahrenheit. To expend minimal energy, we will first use several heat exchangers utilizing the effluent gas from the coal furnace located in the factory. This gas will be sufficiently hot to increase the water temperature a fair amount without requiring any additional energy on our part. To further heat the water to the desired temperature, we believe the reindeer can be useful. By having the reindeer power a treadmill or a turbine, we can simultaneously generate energy for heating and allow them to continue exercising on days other than Christmas.

When the water is heated to an appropriate temperature, we will split it into two separate streams. One stream will pass through a packed-bed reactor filled with peppermint tea leaves which will infuse the passing water. The separate stream for hot cocoa will go into a continuous stir reactor where it will be combined with milk, cocoa, and sugar. Both streams will then pass into the factory itself, where dispensers will be located.

The design can be modified to suit any number of elves, but initially we plan to produce 10 cups of hot chocolate and tea per elf in the factory daily. We can modify the flow of water using pumps located before and after the splitter to suit change in preference. All pipes will be of appropriate gauge and size to match the pressure and temperature of the liquid, and appropriately insulated to weather the cold. The original design will be made slightly larger than required, in case there is a future increase in elves.

Overall, the design should accomplish the task of providing refreshments to the elves. Costs have been minimized by using the reindeer and coal furnaces already located in the plant. Due to an increase in elf efficiency, production will go up and more children will receive presents. We can use the excess coal that was left ungifted to further power the furnace, which will further improve the cost efficiency of our design.