User Testing with Assessors to Develop Universal Rubric Rows for Assessing Engineering Design

Nikita Dawe, University of Toronto

Nikita is a M.A.Sc. candidate in the Department of Mechanical and Industrial Engineering at the University of Toronto. She is completing the Collaborative Program in Engineering Education.

Ms. Lisa Romkey, University of Toronto

Lisa Romkey serves as an Associate Professor, Teaching Stream with the Division of Engineering Science at the University of Toronto. In this position, Lisa plays a central role in the evaluation, design and delivery of a dynamic and complex curriculum, while facilitating the development and implementation of various teaching and learning initiatives.

Lisa is cross-appointed with the Department of Curriculum, Teaching and Learning at OISE/UT, and teaches undergraduate courses in engineering & society, and graduate courses in engineering education.

Lisa completed an Undergraduate Degree in Environmental Science at the University of Guelph, and a Master’s Degree in Curriculum Studies at the University of Toronto. Research interests include teaching and assessment in engineering education.

Dr. Susan McCahan, University of Toronto

Susan McCahan is a Professor in the Department of Mechanical and Industrial Engineering at the University of Toronto. She currently holds the position of Vice Provost, Innovations in Undergraduate Education. She received her B.S. (Mechanical Engineering) from Cornell University, and M.S. and Ph.D. (Mechanical Engineering) from Rensselaer Polytechnic Institute. She is a Fellow of the American Association for the Advancement of Science in recognition of contributions to engineering education has been the recipient of several major teaching and teaching leadership awards including the 3M National Teaching Fellowship and the Medal of Distinction in Engineering Education from Engineers Canada.

Gayle Lesmond, University of Toronto

Gayle Lesmond is a Research Assistant in the Department of Mechanical and Industrial Engineering at the University of Toronto. Her area of specialization is rubric development and testing.
User Testing with Assessors to Develop Universal Rubric Rows for Assessing Engineering Design

Abstract

This paper describes the process of testing and refining modular rubric rows developed for the assessment of engineering design activities. This is one component of a larger project to develop universal analytic rubrics for valid and reliable competency assessment across different academic disciplines and years of study. The project is being undertaken by researchers based in the Faculty of Applied Science and Engineering at the University of Toronto.

From January 2014 to June 2015, we defined and validated indicators (criteria) for engineering design, communication, and teamwork learning outcomes, then created descriptors for each indicator to discriminate between four levels of performance: Fails, Below, Meets, and Exceeds graduate expectations. From this rubric bank, applicable rows can be selected and compiled to produce a rubric tailored to a particular assessment activity. Here we discuss these rubrics within the larger context of learning outcomes assessment tools for engineering design.

We tested draft rubrics in focus group sessions with assessors (teaching assistants and course instructors who assess student work in engineering design). We followed the testing with structured discussions to elicit feedback on the quality and usability of these rubrics, and to investigate how the assessors interpreted the language used in the indicators and descriptors. We asked participants to identify indicators they believed were irrelevant, redundant, or missing from the rubric. We also asked them to identify and discuss indicators and descriptors that were confusing. Finally, we asked them what changes they would recommend and what training materials they would find useful when using rubrics of this design.

We analysed the consistency of assessor ratings to identify problematic rows and used qualitative feedback from follow-up discussions to better understand the issues that assessors had with the rubrics. Three of the six engineering design rubric items we tested showed evidence of rater inconsistency, uncertainty, and indecision. While some rubric rows received similar criticism from most participants, we also identified differences in assessors' rubric design preferences and in how they apply rubrics to evaluate student work. It also emerged that assessors have different conceptions of engineering design and the design process, and are confused when presented with unfamiliar terminology. Based on our interpretations, we identified changes we should make to improve the rubric structure and content.

Our aim is to inform other educators who may be developing tools for the assessment of engineering design by sharing our methodology and discussing the feedback received. Engineering educators could adopt or adapt our user testing methodology to improve the
usability of similar assessment tools. Our discoveries about rubric structure improvements could be explored further to define best practices in the design of universal rubrics. Our next steps include applying what we have learned to refine the rubrics and develop accompanying training materials. The refined rubric rows will be evaluated for inter-rater reliability, trialed in focus groups with undergraduate students, and deployed in academic courses.

Background: Learning Outcomes Assessment and the DARCA Project

There is a need for valid and reliable tools for assessing learning outcomes in engineering education. In the United States the Accreditation Board for Engineering and Technology (ABET) defines curricular requirements for academic programs. Similarly in Canada, the Canadian Engineering Accreditation Board (CEAB) requires that a set of graduate attributes be assessed and reported on for undergraduate engineering program accreditation. Learning outcomes assessment also generates useful data for continuous improvement initiatives within institutions. The University of Toronto is a member of the Higher Education Quality Council of Ontario (HEQCO) Learning Outcomes Assessment Consortium. We are working on the Development of Analytic Rubrics for Competency Assessment (DARCA), a project attempting to produce universal analytic rubrics that authentically assess learning outcomes in engineering design and four other competency areas. The rubrics must be "universal" in the sense that they can be applied to multiple assessment activities to measure performance across different academic disciplines and years of study. In an analytic rubric (as opposed to a holistic rubric) a set of descriptors is provided for each indicator (learning outcome criterion) to define discrete levels of mastery.

Students, administrative staff, accreditation agencies, course instructors, and assessors all play roles in learning outcomes assessment, however their expectations for the rubrics do not entirely overlap. The scope of this paper is limited to feedback from the assessor stakeholder group. Assessors include teaching assistants and course instructors who provide feedback and assign grades to students based on deliverable artifacts as well as performance in the classroom. Assessors may or may not be involved in the design of the assessment activities themselves; some assessors are provided with a marking scheme or a rubric and little further guidance when assessing student work.

This study draws from many areas of literature on learning outcomes assessment and assessment tools as well as engineering design education. The primary focus is on the use of rubrics for assessing engineering design.

Draft Rubric Development

Between September 2014 and June 2015 the DARCA team completed the process of developing draft rubrics. Based on the literature and existing collections of learning outcomes (e.g.
we created an extensive hierarchical list of learning outcomes and their more specific, measurable indicators. This list of outcomes and indicators was validated through expert consultation with engineering design instructors at the university. To produce analytic rubric rows we then formulated descriptions of quality for each indicator to discriminate between four levels of performance: Fails, Below, Meets, and Exceeds graduate expectations. For consistency in description, the levels were defined to mean the following:6:

- Fails: Indicator is not demonstrated (Not Demonstrated) OR complete lack of quality and/or demonstration of a fundamental misunderstanding of the concept (Misconception).
- Below Expectations: Lacks quality; work must be revised significantly for it to be acceptable.
- Meets Expectations: Definition of quality; work is acceptable and demonstrates some degree of mastery.
- Exceeds Expectations: Student goes over and above the standard expectations to produce superior work.

From the extensive list of measurable indicators, specific ones can be selected and compiled to produce a rubric tailored to a particular assessment activity (see Appendix A for an example).

The content of the draft Design Rubric descriptors was produced based on a review of the literature on engineering design competency definitions and assessment with additional input from engineering design instructors. To determine whether the rubric rows are suitable for use by assessors, we developed a focus group research methodology for user testing.

**Focus Group Testing Methodology**

To explore how assessors will use the compiled rubrics and interpret the indicators and descriptors, we conducted focus groups in which participants with experience as teaching assistants and communication instructors for engineering design courses used pre-compiled rubrics to assess sample student deliverables. Between July and October 2015, we conducted three focus groups in which a total of 11 participants assessed interim design report samples from a first year engineering design course. For these sessions we selected 6 Design indicators and 9 Communication indicators to produce an assessment-specific rubric. This paper focuses on the following Design indicators:

- **Outcome D1:** Find and state an engineering design problem
  - Indicator D1B: Accurately state the engineering design problem and summarize key details (interpret a problem statement if provided)
- **Outcome D2:** Gather information to understand an engineering design problem
  - Indicator D2A: Identify stakeholders with interest or influence and accurately describe stakeholder profiles (e.g. characteristics, perspectives, needs)
  - Indicator D2B: Identify and describe engineering design priorities (i.e. Design for X) and/or social and professional concerns relevant to the problem
- Indicator D2C: Extract and integrate information from stakeholders and other appropriate (reliable, diverse, credible) sources to enhance understanding of the problem
- Outcome D3: Frame a problem in engineering design terms
  - Indicator D3B: Document appropriate engineering design requirements using a suitable model (e.g. goals-functions-constraints or objectives-metrics-criteria-constraints)
  - Indicator D3D: Describe the intended engineering design process and provide a plan/timeline that anticipates the tasks and resources required

Indicators D2B, D2C, and D3B were also tested in another session with four participants who assessed sample design proposal assignments for a second year electrical and computer engineering (ECE) course. One of the purposes of this repetition of indicators with a different course assessment piece was to investigate whether the rubric items are generic enough to apply universally across the curriculum.

Focus group sessions began with an overview of the project and an explanation of the agenda and expectations for the session. Participants completed demographic surveys and then were given time to use the rubrics to assess the sample assignments at their own pace. In the first focus group participants were provided with five samples; in subsequent sessions we limited it to two samples to allow more time for discussion. After assessing the assignments participants completed individual written exit surveys followed by a guided group discussion about their experience using the rubrics.

The purpose of the exit survey and discussions was to elicit feedback on the quality and usability of these rubrics, and to investigate how the assessors interpreted the language used in the indicators and descriptors. We believed this would serve as useful qualitative evidence to direct the changes we should make to problematic rubric content. To encourage discussion we asked a series of questions prompting participants to identify indicators they believed were irrelevant or redundant and to discuss language in the indicators and descriptors that was confusing. We also went through the rubric row by row and asked for general comments and feedback. The discussions were recorded for analysis.

**Quantitative Analysis: Rubric Ratings**

After the focus group testing sessions the rubric ratings were analysed by recording the frequency of Fails, Below, Meets, and Exceeds ratings selected by participants in their assessment of the individual assignments. Indicators D2B, D2C, and D3B appeared to be most problematic as ratings were less consistent and assessors did not always select a rating. In some cases participants selected multiple levels or a midpoint. For the quantitative analysis we
assigned half a count to each level (e.g. a Below-Meets midpoint rating was counted as Below: 0.5, Meets: 0.5). We address this assessor tendency in the qualitative discussion section.

For indicator D1B (Accurately state the engineering design problem and summarize key details (interpret a problem statement if provided)) there was some consistency as participants most frequently rated all assignments at the Below level, with a minority rating assignments at the Meets level.

For indicator D2A (Identify stakeholders with interest or influence and accurately describe stakeholder profiles (e.g. characteristics, perspectives, needs)) there was some consistency with participants generally split between two adjacent levels.

Participants did not always select a rating for indicator D2B (Identify and describe engineering design priorities (i.e. Design for X) and/or social and professional concerns relevant to the problem). Ratings were spread across three or four levels indicating poor consistency among assessors.

- DR1: Ratings ranged across three levels from Fails to Meets with one 'no selection'. The majority of participants (5/9) selected Below.
- DR2: Ratings ranged across three levels from Below to Exceeds with one 'no selection'. Participants most frequently selected Meets (3/7).
- DR3: Ratings ranged across three levels from Below to Exceeds with one 'no selection'. Participants most frequently selected Meets (3/9).
- DR4: Ratings ranged across three levels from Below to Exceeds. The majority of participants (4/7) selected Meets.
- DR5: Ratings ranged across all four levels from Fails to Exceeds. Participants most frequently selected Meets (6/11).
- ECE1: Ratings ranged across three levels from Fails to Meets.
- ECE2: Participants selected Below (2.5/4) or Meets (1.5/4).
- ECE3: Ratings ranged across two levels, Meets and Exceeds, with one 'no selection'. Participants most frequently selected Exceeds (2/4).

Participants did not always select a rating for indicator D2C (Extract and integrate information from stakeholders and other appropriate (reliable, diverse, credible) sources to enhance understanding of the problem). Ratings were spread across three or four levels indicating poor consistency among assessors.

- DR1: Ratings ranged across three levels from Fails to Meets with one 'no selection'. The majority of participants (5/9) selected Below.
- DR2: Ratings ranged across three rubric levels from Below to Exceeds with one 'no selection'. Participants most frequently (3/7) selected Below.
- DR3: Ratings ranged across all four levels from Fails to Exceeds with one 'no selection'. Participants most frequently (3/9) selected Meets.
● DR4: Ratings ranged across three rubric levels from Below to Exceeds with one 'no selection'.
● DR5: Ratings ranged across all four levels from Fails to Exceeds with one 'no selection'. Participants most frequently (5/11) selected Meets.
● ECE1: The majority of participants (3/4) selected Meets and the remainder (1/4) selected Below.
● ECE2: The majority of participants (3.5/4) selected Meets.
● ECE3: All participants (4/4) selected Exceeds.

There was very little consistency in the ratings for indicator D3B (*Document appropriate engineering design requirements using a suitable model (e.g. goals-functions-constraints or objectives-metrics-criteria-constraints)*).

- DR1: Ratings ranged across three rubric levels from Fails to Meets. Participants most frequently selected Below (4/9) followed by Meets (3/9).
- DR2: Ratings were polarized; the majority of participants were split between Fails-Misconception (3/7) and Exceeds (3/7).
- DR3: Ratings ranged across all rubric levels from Fails to Exceeds.
- DR4: Ratings ranged across all rubric levels from Fails to Exceeds.
- DR5: Ratings ranged across all rubric levels from Fails to Exceeds. Participants most frequently (4/11) selected Exceeds.
- ECE1: Participants most frequently (2/4) selected Fails-Not Demonstrated.
- ECE2: Participants most frequently (2/4) selected Meets.
- ECE3: 2 participants did not select a rating.

For indicator D3D (*Describe the intended engineering design process and provide a plan/timeline that anticipates the tasks and resources required*), due to a missing appendix in a sample assignment rather than issues related to the design of the rubric, some participants selected Fails-Not Demonstrated while others believed the indicator was demonstrated and assessable. Ratings were otherwise somewhat consistent with participants generally split between two adjacent levels.

**Qualitative Analysis: Feedback on Design Content**

By transcribing and analyzing recordings of the focus group discussions, we identified rubric content that is unclear, ambiguous, or confusing to assessors and synthesized their recommendations for making the rubrics more usable. For the problematic indicators identified above, unfamiliarity with engineering design terms and difficulty distinguishing between levels were common issues elicited from the discussions.

**Indicator D2B:** *Identify and describe engineering design priorities (i.e. Design for X) and/or social and professional concerns relevant to the problem*
Several participants found the term "engineering design priorities" vague or confusing. One guessed that it related to the clarity of the problem stated. Some participants were unfamiliar with the concept of Design for X that was given as an example, while one believed Design for X was not relevant to the document being assessed.

Participants could not tell the difference between Meets and Exceeds. They thought Exceeds would be "going beyond [...] to specify values and stuff".

Indicator D2C: **Extract and integrate information from stakeholders and other appropriate (reliable, diverse, credible) sources to enhance understanding of the problem**

- Several participants found this indicator confusing. They were unsure of its meaning; "what I was marking". One participant said it was hard to judge whether students extracted or integrated information.
- One participant said that assessment of the amount of research was missing.
- Participants agreed that it was hard to distinguish between Meets and Exceeds; they "don't see a lot of difference" and the word skillfully is not measurable or hard to measure. One participant suggested that Exceeds could include evidence of critical thinking and different perspectives instead of probable stakeholders.

Indicator D3B: **Document appropriate engineering design requirements using a suitable model** *(e.g. goals-functions-constraints or objectives-metrics-criteria-constraints)*

- Participants agreed that this is a key indicator for the type of report that was assessed. One noted that engineering requirements are what students "mess up most".
- Participants also agreed that this should be separated into multiple indicators that focus on key concepts. They suggested several different ways of assessing the individual concepts. One participant said multiple indicators might be unnecessary as functions, objectives, and constraints are usually of the same quality; "you understand what needs to go in each, or you don't".
- Some participants couldn't tell the difference between Meets and Exceeds but suggested "the measurable part" would be Exceeds. Another participant suggested that to achieve a rating of Exceeds, students would need to explain the reasoning behind the chosen objectives, provide testable metrics with information and calculations to support values, and consider standards if applicable.

**Refining the Rubric Content**

Given the feedback received in the assessor focus groups, we have identified specific alterations that need to be made to the problematic rubric rows.

Participants were confused by the phrase “design priorities” (indicator D2B) and the example given, Design for X, did not help to clarify what they were supposed to be assessing. The intention of this indicator is to assess whether students are conscious of the real-world
implications of their design problem and defining and prioritizing objectives accordingly. Assessor confusion may be due to the language used in the indicator and its descriptors, or it could possibly be due to lack of familiarity with this type of design thinking. For the purpose of improving the rubric item the phrase “design priorities” needs to be accompanied by an explanation of its intended meaning. Participants interpreted the indicator to mean "prioritized objectives" and “clear distinction” between primary functions and unintended functions. These terms from the discussion could be useful to include as examples that will have meaning to assessors.

Some participants felt they were unable to assess how well students had integrated information and understood the problem (indicator D2C) as they did not observe the research process or have access to the original sources for comparison. This is a recurring issue in the assessment of process-based activities such as engineering design and teamwork, since external assessors only have limited evidence upon which they can make their judgement, and is indicative of a need for ongoing assessor engagement throughout the process. Within the limitations of assessing end deliverables using a rubric, clear descriptors can help to define what assessors should look for as evidence of effective research and immersion in the problem. In this case however, the text in the rubric descriptors overwhelmed assessors and increased their confusion as it was too generic and tied to the process not the product.

Those who felt capable of assessing indicator D2C had different expectations for quality or interpreted the descriptions of quality differently, although participants in the ECE session ended up selecting the same rating in the end anyway. Based on these results, indicator D2C needs to be revised to define more specifically what evidence assessors should be looking for at each rubric level.

Assessors appeared familiar with the concept of an engineering requirements model (indicator D3B) but had different expectations for quality or interpreted the descriptions of quality differently. From the discussion, it appears the inconsistency in the ratings can partially be attributed to the multi-dimensionality of the indicator; it asks for the assessment of multiple concepts in terms of several qualities (e.g. whether objectives are sufficiently measurable, whether functions are appropriate). The indicator needs to be divided into separate indicators, or the descriptors need to more explicitly separate the assessment of the individual concepts.

**Refining the Rubric Structure**

Analysis of the transcripts identified a discussion theme that we had not specifically anticipated: assessor preferences and tendencies relating to the use of a universal analytic rubric. Participants criticized the volume of text they had to process (“it’s just a lot of information”); preferring less verbose descriptors and/or fewer items. Given the quantity of indicators, they struggled to
navigate the rubric as they assessed different elements of an assignment. These comments helped us to identify means of improving the rubric structure.

There is a trade-off between providing enough description to enable reliable assessment, and minimizing the volume of information so assessors are not required to spend too much time reading the rubric. Some participants recommended we “cut down on the description”, especially if they believed descriptors are not useful for giving students meaningful formative feedback: “it’s a lot easier for me to comment on the document [...] versus trying to fit them into such rigid categories”. This is an interesting issue that could be explored further using human factors analyses of rubric information processing. For the purpose of refining our rubrics, we will focus on improving the clarity and conciseness of descriptors where possible.

In terms of the quantity of items, the conflicting rubric functions of assigning grades and providing formative feedback caused issues for some assessors. They found the extensive list of indicators useful for giving specific feedback to students (“many of them were very small [...] you could check very quickly”) but they suggested that the rows might not all be relevant to a calculated grade if the rubric was also used for that purpose. In this situation, instructors could assign different weightings to different rows, or highlight the key indicators that will count towards a calculated grade.

Participants found the rubric difficult to navigate as they did not see any organization in the ordering of the rows. They suggested that indicators be grouped according to the predefined sections of the report (“pick it specific to each section”) or categorized some other intuitive way to expedite the assessment process. Future testing will include trials in which rows are grouped by assignment section when relevant; otherwise rows can be grouped by learning outcome and/or competency.

Another topic of discussion was assessor preferences and tendencies when making markings on rubrics and assignments to communicate their assessments. Participants noted that in some rubric rows checkboxes were used consistently to separate elements, so for example they could select one checkbox under Below and one checkbox under Meets. Some of the rows lacked this consistency and made it difficult for assessors to use the rubric to record and communicate the strengths and weaknesses they observed in the sample assignments. They prefer more flexibility than four defined levels, and described several unique solutions to compensate, for example “I often do a line with arrows going to the different sections”, and “I even mark between categories”. Future iterations of the rubric will make consistent use of checkboxes to separate the elements being assessed within a row.

**Summary of Results and Implications for Engineering Design Assessment**
Three of the six engineering design rubric items we tested showed evidence of rater inconsistency, uncertainty, and indecision. Based on the guided discussion, we have identified means of improving these rows:

- **Indicator D2B** (*Identify and describe engineering design priorities (i.e. Design for X) and/or social and professional concerns relevant to the problem*) needs to be refined so assessors understand the intended meaning of “design priorities”.

- **Indicator D2C** (*Extract and integrate information from stakeholders and other appropriate (reliable, diverse, credible) sources to enhance understanding of the problem*) needs more specific descriptions of quality that tell assessors how to identify well-integrated information and how to gauge the level of problem understanding when they do not have the opportunity to observe the students’ process.

- **Indicator D3B** (*Document appropriate engineering design requirements using a suitable model (e.g. goals-functions-constraints or objectives-metrics-criteria-constraints]*) needs to be reconfigured for unidimensionality of assessment, by separating out the elements of an engineering requirements model into distinct indicators.

Qualitative analysis of the discussions also helped to highlight general preferences and tendencies assessors have when it comes to using rubrics for assessment. These discoveries have implications for the design of the rubric structure itself:

- Descriptors are supposed to ensure more consistent assessment, but assessors can be overloaded with too much text and their performance may decline. In our case description is necessary since the rubrics will be used so widely and benchmarking within or across university departments is unlikely. We must ensure the descriptors are meaningful and succinct.

- The granular approach that breaks the assessment down into many elements is useful for giving students guidance but not all rows will be of equal importance in grading, so instructors may need to assign sensible weightings to the different items. It is also important to group items and sequence the rows logically.

- Assessors feel limited by a rubric, especially when they are forced to select one of four levels. Many prefer to note their feedback on the assignments themselves or mark up the rubric descriptors. Providing multiple checkboxes within a row can enable assessors to give more meaningful feedback while still standardizing records.

Engineering educators could adopt or adapt our user testing methodology to improve the usability of similar assessment tools. The sessions provided valuable feedback although it is necessary to remember that participants are speaking from their own experience as assessors rather than from the perspective of assessment experts, thus it is important to consider other sources when determining which changes to implement. Our discoveries about rubric structure improvements could be explored further to define best practices in the design of universal rubrics.
Future Steps

In combination with evidence from the literature and expert feedback, we will apply the knowledge gained from this study to refine key areas of weakness in the rubric content and structure and to develop training materials for assessors. Our next steps will include a study of inter-rater reliability with a statistically significant number of assessors to determine how reliable these universal analytic rubrics are as assessment tools. User testing with undergraduate students is also needed to determine whether students interpret the rubric language correctly and benefit from its instruction and feedback. The rubric items will be deployed in select engineering courses in the 2016-2017 academic year.

Acknowledgements

This work is supported by the Higher Education Quality Council of Ontario (HEQCO) Contract #12/13-RFP-006-06. For more information on the HEQCO Learning Outcomes Assessment Consortium, please visit http://www.heqco.ca/en-ca/OurPriorities/LearningOutcomes/Pages/Assessment-Consortium.aspx.

References

## Appendix A: Design Rubric Rows Used in Testing

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Not Demonstrated</th>
<th>Fails</th>
<th>Misconception</th>
<th>Below</th>
<th>Meets</th>
<th>Exceeds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D1B</strong> Accurately state the engineering design problem and summarize key details (interpret a problem statement if provided)</td>
<td>Does not state problem or provide details</td>
<td>□ Wrongly states (misinterprets) the engineering design problem □ Provides irrelevant or inaccurate details</td>
<td>Describes the engineering design problem but some details are incomplete, unclear, or inaccurate</td>
<td>Accurately and fully states the engineering design problem and summarizes key details</td>
<td>States the engineering design problem and summarizes key details in a precise and skillful manner</td>
<td></td>
</tr>
<tr>
<td><strong>D2A</strong> Identify stakeholders with interest or influence and accurately describe stakeholder profiles (e.g. characteristics, perspectives, needs)</td>
<td>Does not identify or describe stakeholders</td>
<td>□ Fails to identify interested or influential stakeholders beyond the client or user □ Profile descriptions are misrepresentative or disrespectful</td>
<td>□ Critical gaps in identified stakeholders indicate an underestimation of the project's reach □ Descriptions of stakeholder characteristics, perspectives, and needs are superficial</td>
<td>Identifies expected stakeholders with interest in or influence on the project; accurately represents characteristics, perspectives, and needs in profile descriptions</td>
<td>Evidence of thorough exploration to identify stakeholders with any direct or indirect connection to the project; descriptions are informed and empathetic, capturing characteristics, perspectives, and needs relevant to the engineering design problem</td>
<td></td>
</tr>
<tr>
<td><strong>D2B</strong> Identify and describe engineering design priorities (i.e. Design for X) and/or social and professional concerns relevant to the problem</td>
<td>□ Does not identify engineering design priorities or concerns □ Sufficient descriptions of the design priorities are not provided</td>
<td>□ Identifies irrelevant concerns and/or excludes essentials (e.g. safety, cost) □ Descriptions of the design priorities are erroneous</td>
<td>□ Critical gaps in design priorities or concerns indicate an underestimation of the project's impact □ Descriptions of the design priorities are superficial and not specific to the engineering design problem</td>
<td>□ Identifies expected engineering design priorities or concerns □ Descriptions of the design priorities relate to the specific engineering design problem</td>
<td>Insightful descriptions of both expected elements of engineering design and specific, non-standard elements that suit the problem</td>
<td></td>
</tr>
<tr>
<td><strong>D3B</strong> Document appropriate engineering design requirements using a suitable model (e.g. objectives-metrics-criteria-constraints)</td>
<td>Engineering design requirements not documented</td>
<td>Documents inappropriate requirements or uses an unsuitable model; requirements are inaccurate and/or unmeasurable</td>
<td>Requirements are documented through inconsistent use of a model; requirements lack some precision and measurability</td>
<td>Appropriate requirements are documented using an accepted model; requirements are expressed in precise language and measurable terms</td>
<td>Skillfully integrates information to document comprehensive, precise, and measurable requirements that promote objective judgment of solutions</td>
<td></td>
</tr>
<tr>
<td>Indicator</td>
<td>Not Demonstrated</td>
<td>Fails</td>
<td>Misconception</td>
<td>Below</td>
<td>Meets</td>
<td>Exceeds</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>D3D Describe the intended engineering design process and provide a plan/timeline that anticipates the tasks and resources required</td>
<td>Engineering design process and project plan are not provided</td>
<td>☐ Defines an inappropriate engineering design process given the project context</td>
<td>☐ Engineering design process excludes some key activities and/or contains some problematic components</td>
<td>☐ Plan is not feasible but excludes some critical tasks or resources</td>
<td>☐ Defines an appropriate engineering design process based on an accepted model</td>
<td>☐ Integrates academic and industry practices to define an engineering design process that is tailored to the problem</td>
</tr>
<tr>
<td>D2C Extract and integrate information from stakeholders and other appropriate (reliable, diverse, credible) sources to enhance understanding of the problem</td>
<td>No information extracted</td>
<td>Information is inappropriate and does not enhance understanding of the problem</td>
<td>Sources are minimally appropriate; understanding is insufficient and with gaps, and will limit problem framing</td>
<td>Uses multiple strategies to extract and integrate information from stakeholders and other appropriate sources; understanding is sufficiently enhanced for problem framing</td>
<td>Skillfully extracts, evaluates, and integrates information from a variety of authoritative sources; understanding is greatly enhanced, enabling informed problem framing</td>
<td></td>
</tr>
</tbody>
</table>

A full set of the rubric rows created is available publically: http://sites.google.com/site/uoftlearningoutcomesproject