SURVEYS FOR SIMULTANEOUS DIRECT AND INDIRECT ASSESSMENT

Derek Wright, Emma Markoff, Jason Grove, Rania Al-Hammoud, Andrew Milne, and Mehrdad Pirnia
Faculty of Engineering, University of Waterloo
derek.wright@uwaterloo.ca

Abstract – As the Canadian Engineering Accreditation Board moves toward outcomes-based assessment, engineering institutions are employing a variety of assessment tools and methods. Surveys are usually thought of as indirect measures. However, survey tools can easily provide direct measures as well, and can therefore simultaneously present direct and indirect assessments to students. This paper examines the benefits and possibilities of exploiting such simultaneous direct and indirect assessment.

Keywords: Survey, Direct Assessment, Indirect Assessment, Lifelong Learning.

1. INTRODUCTION

Canadian engineering institutions are employing a variety of assessment tools and methods to meet the Canadian Engineering Accreditation Board (CEAB) outcomes-based assessment criteria. These can be grouped into direct and indirect depending on the tool and its use. For example, a professor using a rubric to assess a student’s performance in a learning activity is a direct measure, and a focus group of students discussing the impact of that same learning activity is an indirect measure of their performance.

Surveys are usually thought of as indirect measures. At the University of Waterloo (UW), for example, we ask students to rate their performance as Beginning, Developing, Accomplished, or Exemplary for each program-level indicator statement and according to a rubric that defines these levels. These survey questions are clearly indirect measures and represent a student’s perception of their abilities instead of their actual abilities.

However, survey tools can easily provide direct measures as well. Multiple choice questions, freeform text, rank order, item grouping and ordering, and peer assessment are all common capabilities of modern online survey tools. These features can be used to test student performance directly. For example, concept inventories can be implemented with multiple choice questions, students can correctly order process steps, and text boxes for reflection questions can all be easily presented to students.

Therefore, survey tools can simultaneously present direct and indirect assessments to students. Aside from the logistical benefit of collecting two assessment types at once, two other notable benefits arise stemming from the fact that the results can be correlated per individual. The first is that indirect survey data is often considered the “lowest form of data,” and such a correlation, if present, serves to validate the student self-assessment, both within the specific survey and more broadly as a useful method. A second, less obvious benefit is that the correlation itself between the self- and direct assessment could be viewed as a proxy measure for certain Lifelong Learning indicators.

The twelfth CEAB graduate attribute, Lifelong Learning, is defined as “an ability to identify and address their own education needs in a changing world in ways sufficient to maintain their competence and to allow them to contribute to the advancement of knowledge”. UW has defined three indicator statements for this graduate attribute, the first of which states “Graduates should be able to identify gaps in their knowledge, skills and abilities” (Bloom’s Taxonomy verb in bold). The hypothesis is that a high correlation between self- and direct assessment points to high self-awareness of one’s abilities, whereas a low correlation points to low self-awareness. To this end, we have been examining such correlations for use as indirect measures of this performance indicator.

The remainder of this paper is as follows: In Section 2 we will present and discuss the gathering and analysis of simultaneous direct and indirect assessments in three engineering courses at UW, and in Section 3 we will discuss these results, propose future research, and draw conclusions.

2. TRIALS

Three cohorts of students were assessed in three different Electrical and Computer Engineering (ECE) courses: 1) ECE 498B – Engineering Design Project; 2) ECE 140 – Linear Circuits; and, 3) ECE 200B – Electrical and Computer Engineering Practice. In all cases, students were asked to perform self-assessments based directly on
indicator statements and then were asked to assess their teammates on those same statements in the case of ECE 498B, or to solve some problem that resulted in a direct assessment of their performance. The students used the rubric in Table 1 for self and peer assessments.

The ordinal assessment levels (Beginner, Developing, Accomplished, Exemplary) were converted to cardinal values (0 through 3, respectively) for the purpose of analysis. This mapping may be problematic in that the ordinal levels may not fall neatly on a linear cardinal scale, but it is reasonable to assume that the levels in Table 1 do represent an increasing scale. The subsequent analyses use this mapping where noted and with this caveat in mind.

2.1. ECE 498B

The goal of this analysis was to understand the congruency between self-ratings and others’ ratings (self-other rating agreement, SOA [1-2]), which may indicate whether students can identify gaps in their knowledge and learning. This course is the second half of the ECE capstone design project and occurs in the students’ final academic term, and was chosen because of the large population of enrolled students (263 students). The projects are completed in teams of three to five students, all of whom are usually from ECE. Teams are formed in the spring 4A term as part of the related ECE 498A course and last until the project’s completion in the following winter term, meaning that groups persist for approximately 11 months. This duration of exposure to one another plus the stressful conditions under which the exposure occurs means that teams should have a good sense of each other’s strengths and weaknesses.

The five questions posed to students with regard to themselves and their group members are listed in Table 2. The questions were reworded for group members so that “you” was replaced with the group member’s first name. Respondents chose answers from the rubric in Table 1, with an additional “Not sure” option presented during group member assessment. The survey was optional for students to complete, and the final sample size was 95 students. The low sample size makes it unlikely that a statistically significant dataset was extracted from this survey. However, this does not eliminate the value of any unusual patterns that may inform future research.

![RMS Self-Other Difference vs. Self-assessment of Lifelong Learning](image)

Fig. 1. RMS of self-other difference vs. self-assessment of Lifelong Learning (N = 95).

The SOA was determined by calculating the root mean square (RMS) of the per-question differences between each self-assessment and the average assessment of that same question by team members. This was then compared to each person’s self-assessment of Lifelong Learning. One would expect that as true performance in the Lifelong Learning indicator increase the SOA also increases. So, if

<table>
<thead>
<tr>
<th>Levels</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginner</td>
<td>I have limited knowledge and require frequent assistance or guidance to perform related tasks.</td>
</tr>
<tr>
<td>Developing</td>
<td>I have adequate understanding / knowledge in performing related tasks. I require minimal supervision to perform related tasks.</td>
</tr>
<tr>
<td>Accomplished</td>
<td>I have a higher level of understanding and usually do not require assistance to perform related tasks.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>My knowledge of the subject is outstanding. I am often considered a “go to” person for this topic.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Management</td>
<td>At what level can you <strong>apply</strong> project management techniques in engineering projects, with attention to risk, change, and limitations? At what level can you <strong>create</strong> a substantial engineering solution</td>
</tr>
<tr>
<td>Design Process</td>
<td>Through iterative design while considering requirements and constraints? To what degree can you <strong>perform</strong> constructively as a team member?</td>
</tr>
<tr>
<td>Teamwork</td>
<td>Qualities include participation, cooperation, and flexibility. To what degree can you effectively <strong>manage</strong> your time based on priorities?</td>
</tr>
<tr>
<td>Time Management</td>
<td>To what degree can you effectively <strong>identify</strong> gaps in your knowledge, skills, and abilities?</td>
</tr>
</tbody>
</table>

**Table 1. Self and peer assessment rubric.**

**Table 2. ECE 498B assessment questions.**
The self-assessment of the Lifelong Learning indicator correlates positively with true ability on this indicator, then the RMS differences should decrease with increasing self-assessment. This trend does appear to be weak but present in Fig. 1.

2.2. ECE 140

This course teaches introductory circuits to 1A ECE students and was chosen because it had a large number of enrolled students (552 students). Several course-level indicators were assessed, but Error! Reference source not found. shows only those where direct and indirect assessments were paired. They both map to ECE program-level indicator 1a: “Apply knowledge of the fundamental science and mathematics relevant to EE or CE”. Students were given the self-assessment rubric presented in Table 1 and asked to rate themselves for each course-level performance indicator.

Performance indicators a) and f) were also directly assessed by using a concept inventory from the Appraisal System for Superior Engineering Education Evaluation-instrument Sharing and Scholarship (ASSESS) search engine for assessment instruments [3]. Specifically, questions from the thesis entitled “Enactive Modeling as a Catalyst for Conceptual Understanding” were chosen [4].

The three questions used to measure performance indicator a) are shown in Fig. 3 with correct answers in bold. Students received 1/3 mark for each correct answer, providing a quantitative assessment of student performance. Performance indicator f) was assessed using the question shown in Fig. 2, where a full mark was given for the correct answer.

Table 3. ECE 140 program- and course-level indicators.

<table>
<thead>
<tr>
<th>Program-Level Indicator</th>
<th>Course-Level Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>a) I can analyze linear circuits by applying Ohm's Law and Kirchoff's Voltage and Current Laws to find node voltages and mesh currents</td>
</tr>
<tr>
<td></td>
<td>f) I can analyze sinusoidal steady-state circuits and determine AC power</td>
</tr>
</tbody>
</table>

The headlights on a toy car are wired as shown below. Assume that the headlights are identical. For the next three questions, choose the statements that best describe this arrangement.

- Q4) When the switch is closed,
  - Headlight A will be brighter than headlight B.
  - Headlight B will be brighter than headlight A.
  - Both headlights will be equally bright.
  - Neither headlight will light at all.

- Q5) If headlight A burns out,
  - Lamp B burns more brightly than it did before the burnout.
  - Lamp B burns just as brightly as it did before the burnout.
  - Lamp B lights but burns less brightly than it did before the burnout.
  - Lamp B does not light.

- Q6) An engineer wants to redesign the car in such a way that the headlights shine more brightly. Which of the following designs will make the headlights shine more brightly?
  - Reverse the way the battery is connected.
  - Move the battery so that it is in between the two headlights.
  - Rewire the circuit so that the headlights are in parallel with the battery.
  - The brightness cannot be increased without changing or adding batteries or using different headlight bulbs.

Fig. 2. Direct assessment questions for ECE 140 course-level indicator f)

Q7) When she cranks the frequency to 120 Hz she finds that the ammeter reading is now:
- 1 A
- 2 A
- 0.5 A
- 0 A

An engineer connects a capacitor to a variable-frequency signal generator and measures the current with an ammeter. At 60 Hz, she finds that the ammeter reading is about 1 A.
The results in Fig. 4 for performance indicators a) and f) show that students who rated themselves higher did better than those who rated themselves lower. For performance indicator a) (Q4 through Q6), there is little difference in direct performance for Developing through Exemplary, resulting in a low correlation ($R^2 = 0.6651$) between self- and direct-assessment. The results for performance indicator f) show a good correlation ($R^2 = 0.9392$) between self- and direct-assessment, however, only 48% of students who rated themselves as Exemplary chose the correct answer. So perhaps the students have a good relative sense of their abilities, but a poor sense in absolute terms.

![Fig. 4. Direct vs. self-assessment for ECE 140 performance indicators a) and f) (N = 119)](image)

### 2.3. ECE 200B

This course was chosen because one of the authors (D. Wright) was the instructor, and the nature of the course makes it useful as an “assessment testbed”. The course was modified so that students stepped through a printed circuit board (PCB) design workflow, which provided an opportunity to assess a variety of outcomes and performance indicators. The relevant program level outcomes and corresponding performance indicators for ECE 200B are listed in Table 4. These performance indicators represent the expected performance abilities of an ECE 200B graduate. Indicators a) and f) were also directly assessed by using questions developed by Derek Wright.

#### Table 4. ECE 200B program- and course-level indicators.

<table>
<thead>
<tr>
<th>Program-Level Indicator</th>
<th>Course-Level Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>4b</td>
<td>a) I can remember and describe the workflow for designing and fabricating PCBs</td>
</tr>
<tr>
<td>1b</td>
<td>f) I can assemble electronics parts to a fabricated PCB</td>
</tr>
</tbody>
</table>

Q5) Please drag and drop these steps into the right order for a typical hardware design and fabrication workflow. Normally there is a lot of iteration between stages, but put them in the approximate order in which they occur.

- Create system diagram
- Choose major components
- Create schematic
- Create layout
- Export files for fabrication and send to fab
- Order parts
- Assemble board (solder)
- Test and debug

![Fig. 5. Direct assessment questions for ECE 200B course-level indicator a)](image)

Q6) Choose whether each image represents a **good** or **bad** solder joint or method.

![Fig. 6. Direct assessment questions for ECE 200B course-level indicator f).](image)
Fig. 5 shows the question (Q5) used to assess performance indicator a). Using a feature of the Qualtrics survey platform, students were shown the elements of a hardware workflow presented in a random order. The students needed to drag and drop the items into the correct order, and deviation from this correct order was calculated via a normalized root mean squared (RMS) error method, thus providing a quantitative assessment of student performance. The RMS method also provides some leeway in the ordering of the workflow since in reality certain items could be exchanged. Fig. 6 shows the question used to measure performance indicator f) with the correct answers preselected. Students received 1/6 of a mark per correct answer to determine their final score on this question.

![Bar chart showing Direct vs. self-assessment for ECE 200B performance indicators a) and f) (N = 184)](chart.png)

Fig. 7. Direct vs. self-assessment for ECE 200B performance indicators a) and f) (N = 184)

The results in Fig. 7 show that students who rated themselves higher did not necessarily do better than those who rated themselves lower. For performance indicators a) and f) (Q5 and Q6, respectively), there is poor correlation between direct- and self-assessment. Since the course material is quite new to the students, perhaps they lack the ability to judge their understanding and skills appropriately.

3. DISCUSSION AND CONCLUSIONS

Though the relationships examined in this paper were not strong, it seems that characterizing the correlations between direct and self-assessments may provide an indirect measure of a Lifelong Learning indicator. Further study is required to gather enough data to examine these trends at levels of statistical significance. Regardless of the ultimate outcome, it is informative to place direct and self-assessed data side by side to see how accurately students can gauge their capabilities.

References


