NETWORK-BASED APPROACH TO ASSESSMENT OF COGNITIVE SKILLS

Natalie Simper, Nerissa Mulligan, and Brian Frank
Queen’s University
brian.frank@queensu.ca

Abstract - Cognitive Assessment Redesign (CAR) project is an institution-wide, network-based approach to the development of cognitive skills in undergraduate education. This project aims to encourage first and fourth-year instructors to align skill development through the design of course assessments, to enhance cognitive skill acquisition and provide a measurement of learning. The learning outcomes for the project are framed and operationalized using the language and dimensions from the Valid Assessment of Learning in Undergraduate Education (VALUE) rubrics. An assessment redesign network was created, matching assessment facilitators who have disciplinary and educational expertise with instructors to develop authentic assessments of student learning. One of the goals of the network is to encourage sustained participation and collaboration, and to build progress in teaching and learning throughout the institution. The project also includes a standardized test for comparison to course assessment outcomes. Testing at the fourth-year level has been dependent on the use of incentives for student participation. Although recruiting instructors from the faculty of Engineering and Applied Science was initially a challenge, course instructors have reported various successes stemming from participation in the project.

Keywords: networked-based change process; cognitive skills; assessment of learning

1. INTRODUCTION

Well-designed tasks can provide feedback to the instructor about student performance and evidence about long-term student learning, as well as be a learning experience for students. Crafting such tasks can be both time-consuming and intellectually challenging for an instructor due to the need to carefully align the task to both course and program outcomes. Producing change in teaching and learning practices within an institution requires sustained participation, collaboration, and support for participating faculty members [1]. This can come in the form of a top-down organizational change initiative [2], but social networks are particularly critical in decentralized institutions where faculty have tight professional connections to peers at other institutions [3], [4].

Connecting instructors with a network of peers working on similar initiatives provides professional and personal support for engaging in a challenging change process.

In 2017 Queen’s University started the Cognitive Assessment Redesign (CAR) project, an institution-wide, network-based research project focusing on the development and assessment of cognitive skills (critical thinking, creative thinking, and problem-solving) in undergraduate education. The project used a triangulated approach to demonstrate the “value-add” of student learning in undergraduate education at Queen’s. Cognitive skills of first and fourth year students were assessed using the following measures:

- Course assessment rubrics: Developed by the course instructor, in collaboration with the research group
- Standardized rubrics: Valid Assessment of Learning in Undergraduate Education (VALUE) rubrics, developed by the Association of American Colleges and Universities
- Standardized test: HEIghten, developed by the Educational Testing Service (ETS).

The project involves incentivized recruitment of first and fourth year course instructors interested in purposefully aligning skill development through the design of course assessments tailored to cognitive skill acquisition. A central network was created to financially support faculty representatives and encourage interaction between network members for capacity building. The initial goals of the CAR project are to:

1. Build departmental capacity; recognize faculty leaders within departments.
2. Support instructors to develop course-based assessments (in first and fourth-year courses).
3. Validate the course-based outcomes by rating a sample of assignments using the VALUE rubrics and with selective testing using the “HEIghten” test.

Continuing from these, the team plan to develop an institutional guide, then expand the initiative within and possibly beyond Queen’s University.
2. METHODOLOGY

The CAR Project was designed as a systems-level change initiative working to establish a shared vision at multiple levels within the organization [5]. The principal, provost, deans, and associate deans provided input about the goals, people, and approaches in the project. The research team felt it was important to target the largest disciplinary sectors at the university, including Science, Social Science, Humanities, and Applied Science (Engineering). Within each sector, the team sought expressions of interest from an instructor of a large first year course and four instructors of upper year courses in order to provide evidence about student learning over the duration of the program. In order to participate in the project, interested instructors submitted an expression of interest for course assessment structures that apply cognitive assessment to the student learning experience. The deadline for fall term courses was July 2017, and the deadline for winter term and full-year courses was October 2017. Successful applicants received a mini-grant (of up to $5,000) to cover the costs of the proposed assessment intervention, or for scholarly activities, either related to the evaluation of the course assessment or activities that directly contribute to the improvement of student learning.

Maintaining a network however, requires leadership and ongoing commitment, therefore, institutional support plays a key role. In the academic setting, such networks are sometimes referred to as “communities of practice” [6]. Once instructors were approved to participate in the project, they were linked with educational professionals, as part of the assessment redesign network. Additional aspects of the network included regular consultation and planning among participating faculties and employing a research coordinator/manager. The research manager was essential for communication, following protocol, management and reporting for the project, as well as key to the success of faculty engagement, data collection and sustained change management.

The network (see Error! Reference source not found.) is comprised of the CAR members (course instructors), their assessment facilitator, and network support personnel (Centre for Teaching and Learning). Embedded disciplinary expertise is a promising approach to engaging instructors in sustained change initiatives [7]. The assessment facilitators need to have disciplinary as well as pedagogical expertise related to assessment, such that they are able to translate general educational principles into a disciplinary context.
2.1. VALUE rubrics

The VALUE Rubrics are broad, discipline-neutral descriptions of selected essential learning outcomes of undergraduate education from the Liberal Education America’s Promise (LEAP) initiative. There are four levels of performance criteria, from the benchmark level of a student entering university to the capstone level of a student who has just completed their undergraduate experience. Of the sixteen VALUE rubrics created for different skills, the CAR project used the ones pertaining to critical thinking, creative thinking, and problem-solving. These three rubrics are used to enable the comparability of skills across different disciplines.

The learning outcomes for the project are framed and operationalized using the language and dimensions from the VALUE rubrics (see Figure 2). Authentic tasks (assignments) are designed by the instructor with support from the assessment facilitator. The instructors or Teaching Assistants use assignment rubrics for course assessment; they align to the course learning outcomes, but observe (as closely as possible), the language and dimensions from the VALUE rubrics. Where possible, students complete a standardized test, as described below. Course artifacts from consenting students are provided to the research team and scored using VALUE rubrics. Course-based marks are investigated through a triangulated validation approach (see Figure 3).

2.2. Authentic Tasks

The alignment of course assignments with desired assessment criteria often requires the development of tasks that take time and expertise to develop and commitment from those who teach the students is recommended [8]. It is proposed that cognitive skills can be developed through performance-based assessments or authentic tasks that may resemble analysis and research projects, design projects, investigations or structured inquiries [9]. Authentic tasks are the vehicle used to measure students’ ability to apply knowledge or skills as they would in the real world. The problems presented in authentic tasks ask students to meaningfully apply concepts to solve problems, use evidence to recommend solutions, or creatively develop new approaches or methods. They regularly require consideration of context, purpose, outcomes and implications.

Course-based assessments have greater face validity and relevance for the student than a standardized test. However, using course based assessment as a comparable tool to evaluate learning within and across an institution has potential challenges with regard to reliability. Tailoring the assessment to the student’s program of study may lead to a more valid assessment of their learning [9], but differences in assignment type can account for up to 77% of error variance [10], impacting the reliability of the results. It is suggested that an assignment tasks be carefully designed to align with the intended outcomes, and that faculty be directly involved in assessing the quality of student learning [11]. Earlier research by the CAR project team found significant and problematic misalignment between existing course assignments and the assessment criteria used to assess long-term student development [12], which resulted in poor correlations between course grades and scores on program-wide outcomes. This previous study found that many of the course assessment rubrics did not align with equivalent dimensions on the VALUE rubrics.

As such, students were not guided to provide responses that demonstrated the breadth of the critical thinking and problem-solving criteria. For the current study it was important to consider both the nature of the assignment guidelines and the dimensions of the course rubric to ensure valid alignment to the assessment dimensions.

Figure 2. VALUE rubric framework for development of cognitive skills
To support the alignment of assessment dimensions, the research team developed a rubric-building tool for assessment of cognitive skills. The tool, named Building Assessment Scaffolds for Intellectual Cognitive Skills (BASICS) [14] has five-steps prompting instructors to define and articulate the intended learning outcomes for their students, with specific criteria for particular levels of attainment. The assessment facilitators used the BASICS tool as part of the consultation process for redesigning course rubrics.

2.3. Standardized Testing

While program-wide assessment of authentic tasks was the key focus of this project, variation between tasks presents a challenge to the reliability [10]. In order to evaluate the long-term reliability of this approach, and to measure student learning over time, the project also used a standardized instrument. Key experts recommend implementation of assessment approaches that align with course outcomes and that yield actionable evidence to ensure students actively participate in the standardized testing process [13]. The research team recognize the strengths, drawbacks and limitations inherent in standardized testing. For example, testing imposed as a mandatory accountability measure has been shown to have little or no effect on student learning [14]. In addition, standardized tests that are implemented separately from the course experience are susceptible to a decrease in student motivation. If students do not feel motivated to put effort into a test, the resulting score may lack reliability, challenging the validity of inter-institutional comparison [15]. Effort levels in standardized testing conducted in the team’s previous research [12], presented a significant threat to test reliability.

Careful planning was undertaken to mitigate the above factors, as the project team elected to use the HEIghten test, developed by the Educational Testing Service (ETS). It is a 45-minute online critical thinking test. Students can take the test on-site, or un-proctored as a take home activity. It is comprised of a series of questions related to stimulus reflecting real-world issues. It requires students to draw conclusions about what is required/permit by those conditions. Including a log-on procedure, tutorial and sample test, the test can take 60 minutes in total. The test also provides students with an individual statement of achievement displaying their score in comparison to the international average.

3. IMPLEMENTATION

Although there were three areas of focus for cognitive skill development (critical thinking, creative thinking or problem-solving), the majority of instructors focused their assessment redesign efforts around critical thinking and problem-solving.

The mini-grants provided to instructors were used for various purposes, such as supporting an additional TA in the course, supporting a new peer feedback process, implementing a field trip, conference travel to talk about the work, student poster printing, food for a new student presentation session connected to the revised assessment, a guest lecturer, a local student conference, an expert facilitator, as well as a research assistant.

The mini-grants also provided the opportunity for research personnel to conduct marker training sessions in some of the participating courses. These were designed to support the implementation of the new course assignment rubric. The first phase of the marker training sessions involved the instructors and their Teaching Assistants (TAs) reading through and discussing each element of the rubric, specifically, the terminology used, and the hierarchical differences between levels. Follow up meetings were conducted after the independent marking of a selection of work samples. During marking, comments were recorded to provide rationale for the discussion of work criteria. The comparison and discussion of differences in interpretation were used to building marking consistency. These sessions have also proved to be valuable to instructors in validating the way they articulated their student learning outcomes.
3.1. Working with instructors and developing rubrics

The assessment facilitators have had a great deal of success in working with the instructors involved in the project, but it has been challenging to redevelop course assessments. At the half way point in the project implementation, the main challenges facing the assessment facilitators were:

- Figuring out what the instructor was “looking for”, and helping instructors articulate what it was they wanted students to be able to do.
- Rubric generation- In many cases, involvement in the project meant drafting a completely new assignment. One that reflected authentic learning for students, getting at current issues and engaging students in meaningful learning experiences. These new assignments and their associated assessment rubrics were time consuming to develop, and required ongoing refinement.
- Rubric review (aligning course rubrics with the VALUE dimensions while respecting disciplinary differences).

The assessment facilitators worked with the course instructors to review existing practice and past assignments, and asked questions like “What are the deficits? What aren’t they doing?”. The VALUE rubrics provided the framework for development and refinements to course assessments. For example, the VALUE criteria for critical thinking, conveys the need for students to identify their assumptions, then extending from that, to questions their own and others assumptions. Many of the existing assignments asked students to generate a hypothesis and draw conclusions, but not to specifically articulate their tacit or implicit assumptions. It wasn’t just the assignment description and rubric that were redesigned. The learning experience for students was often adapted to achieve the assessment goals. In some cases that involved redesigning one of the labs, or engaging students through student-centred inquiry. In some of the courses, visiting speakers provided authentic connections to real-world problems, whereas another course instructors opted to run a research poster session to prompt students to think about their work from alternative perspectives. In short, the strategy for engaging students with an authentic learning task was different for each of the courses involved.

The instructors provided written feedback, describing their experiences. Investigation of the responses found that there were concerns over the onerous nature of assessment, but the comments were almost unanimously positive, such as “TAs were wary prior to this lab, as many felt it was going to be too easy. The student feedback, and the TA experience, showed that this lab was appropriately targeted for many students (e.g., those who are mid-range)” (CAR course instructor #1). Another instructor commented “TAs reported multiple students coming and thanking them for taking the time to go through critical thinking with respect to research. The TAs were quite excited to see the outcome. I had never had this happen for a research methods topic, so it was great to see a positive buzz!” (CAR course instructor #2).

3.2. Implementing testing

Completion rates for testing were dependent on the approach the instructor took to embedding the test in the course. Instructors that volunteered to have their class involved in testing either offered a nominal attendance grade (typically 1 – 5%), or included the result of the assessment as a nominal percentage of the course grade. The majority of first-year instructors were open to incorporating standardized testing in their course, with 25% opting for in-class testing, 50% of instructors choosing to embed the test outside of class time, but 25% opted not to include the test as part of the course. It was more challenging to test in the fourth-year courses. Attendance in first-year courses is fairly consistent and the subject matter general enough that a standardized test can be incorporated, but fourth-year courses are often technical electives, making it more difficult to incorporate a test that does is not directly linked to the subject matter. Fourth-year courses with unused lab or tutorial sessions, where students were provided food and offered a small percentage of their grade as an attendance incentive had the highest take-up rate. Where there was no class time available for testing, students were invited to complete the unproctored version of the test, with a giftcard raffle incentive.

4. PRELIMINARY CONCLUSIONS

4.1. Building departmental capacity

Incentivizing instructor involvement differed between disciplines. The majority of the early expressions of interest in the project were from Humanities departments, followed by Social Sciences. Most instructors from Science and Engineering departments were recruited by assessment facilitators, as the $5000 mini-grant was less attractive in the disciplines that tend to receive larger funding grants.

Each of the instructors presented their assessment initiatives as “lightning talks” to their peers and wider network in the fall and winter terms. These talks prompted discussion between the members of the various learning areas involved. There was unanimous praise for the assessment facilitators, and discussion of the next steps for further refinement of assessment processes.
4.2. Support instructors to develop course-based assessments

The main successes have revolved around increased awareness, and approach tendencies toward assessing critical thinking. As the VALUE rubric language was adopted, it gave instructors the opportunity and confidence to engage in more informed conversations with their students. By working with the project personnel, in particular the assessment facilitator, allowed instructors to clearly articulate their intended outcomes for students. Some instructors reported that they had fewer questions from students when the course rubrics were explicit, and shared with students ahead of submitting their assignments. Instructors have found the support they received to be very beneficial.

4.3. Validate the course-based outcomes

Generalizing learning across institutions is limited with VALUE data. The general nature of the VALUE rubric is one of its biggest advantages, as the rubrics can be applied to many programs. However, this is also one of the disadvantages, since it is not specifically tailored to any one program. The research team are continuing with their quantitative analysis of results, comparing standardized test results to the course assessment and VALUE scoring. Included in the analytical design is the plan to compare scores from multiple assignments from the same individual, completed in different courses. This will help the research team to evaluate the cognitive skill development in a program and across the institution.

Acknowledgements

This project is supported by the Higher Education Quality Council of Ontario (HEQCO), through the Learning Outcomes Assessment Consortium.

References