EOAS CONCEPT INVENTORY OUTCOMES, SUMMARY OF WORK DONE AND MATERIALS ACCUMULATED AS OF AUGUST 2012.

Skylight project title: "Optimizing the use of concept inventory testing to benefit student learning in EOS undergraduate programs"; **Summer 2012 (July 5th-August 30th)**

Contact: Alison Jolley (alisonjolley@gmail.com); Francis Jones (fjones@eos.ubc.ca)

CONTENTS

BACKGROUND

What is a "concept inventory"?

There is no one unique definition in the literature for "concept inventory". For our purposes, we use it to describe an assessment tool intended to help measure some aspect of teaching or learning. It is focused on key concepts as determined by the relevant instructors (ideally through learning goals). This does not mean that it doesn't include any recall questions at all, but that there are few. The majority of questions are meant to assess higher level thinking, and often have undergone a fairly lengthy and rigorous validation to ensure this.

Project objectives

- Gather, organize, and prepare results of CI's produced in EOAS since 2007.
- Assess individual questions.
- Develop strategies for incorporating CIs and results into EOAS through recommendations, templates and frameworks.

Intended audience

- TIC committee – feedback will help them set education initiative priorities in EOAS.
- Instructors of relevant courses – recommendations should be communicated to them so that they can consider implementing them with STLF assistance.
- STLFs – discuss this with CWSEI people, in order to disseminate experiences to directly relevant audience.

WHAT HAS BEEN DONE

WHAT CONCEPT TESTS HAVE BEEN USED IN EOAS?

Table 1 is a list of inventories used within the Department and various details about their usage – administration timeline, if/how the data were used, what level of validation they've undergone and who to contact if you have further questions (the "Summary of CIs" Excel file contains even more detail).

EXAMPLES OF HOW THE DATA WERE USED

Table 1 includes columns indicating whether CI was used for research, feedback to instructors, feedback to students, and/or for adjusting content or strategies in the course. Examples of these uses are given in Table1 (end of document):

1) 14 of the 19 inventories are being used for **instructor feedback**, mainly to determine specific student backgrounds/experiences that aren't necessarily consistent with pre-requisites. To a lesser extent, pre-post shifts are being utilized for driving pedagogical change.
    a) E.g., EOSC211 (contact: Rich or Catherine), where students enter the course with varying abilities regarding computers and computer programming. The pre-data are particularly important in informing instructors about the incoming population, with both background-type questions and questions asking about evaluating expressions in computing language. The post-data tell them about how students' level of comfort and confidence with computing concepts has changed, as well as their ability to evaluate programming expressions.
2) 12 of the 19 have been used for **research**, to investigate some aspect of student thinking or conceptual knowledge.
    a) E.g., EOSC212 (contact: Francis), where students complete a pre/post activity centered on model based reasoning. The data from this have been used in research regarding development of scientific thinking skills.
3) 9 of the 19 have been used for **student feedback**, which is for the most part generic, and may be referred to all at once or at relevant points during the term. Some use the instrument as a tool for student review and give instant feedback (via Connect or iClickers).
    a) E.g., EOSC 372/373 (contact: Susan or Maite), where students answer questions via iClicker on the last day of 372 and are used to review key concepts and guide studies. These questions are given again at the beginning of 373 in the following term.
4) 5 of the 19 have been used for **course feedback**, to design or change lectures/activities. This seems to be pretty rare, possibly because it requires a higher commitment in terms of time.
    a) E.g., EOSC340 (contact: Sara), where worksheets have been developed to target concepts from specific questions on the pre/post.
5) 3 of the 19 are/have been used for **all of the above**, and the most consistently administered/longest running ones are used for several.

EXTERNAL CONNECTIONS

1) Geoscience Concept Inventory (GCI: http://geoscienceconceptinventory.wikispaces.com/)
   a) Development of this initiative is now based on community discussion
   b) 83 members but no one has contributed in last 8 months+, so nothing is really moving forward
   c) Can submit questions with a short explanation of where misconceptions were observed, but they may rewrite them
   d) Then would be validated by GCI team and community
   e) Doesn't seem to be worth a significant investment, but could contribute a small subset of carefully selected questions just to test it out
2) SERC (http://serc.carleton.edu/index.html)
   a) Large question and activity bank but no details given re: validation
   b) Good starting point for ideas when developing future inventories
   c) Can submit assessment tools to them
   d) GTAT was submitted to SERC – see http://serc.carleton.edu/NAGTWorkshops/time/assessments/60771.html
3) University of Canterbury, Christchurch, NZ (Ben Kennedy, ben.kennedy@canterbury.ac.nz; Erik Brogt, erik.brogt@canterbury.ac.nz)
   a) Attempted to get in touch twice, haven't heard back.
   b) Four inventories (validated with experts, not students): introductory geology, environmental geohazards, minerals and ores, and magmatic systems
   c) Final report has been produced for Ako (funding agency; http://akoaotearoa.ac.nz/projects/transforming-tertiary-science-education), currently undergoing peer review but should eventually be up online
   d) Not sure exactly what the future plans are in terms of educational measurement, but are creating a formal geoscience education research group within the department

RECOMMENDATIONS

These are ideas that arose as data were compiled and processed. The list is neither exhaustive, nor final. It is meant to spark interest where possible.

HIGH PRIORITY INVENTORIES

These are defined as high priority mainly because of a combination of a high level of validation, identified area of need, and/or large dataset already in existence. Or because recommended actions are "easy" (e.g., 211).

1) Min/Pet (220, 221, 321, 322): largely validated, with minor work to be done refining questions.
   a) There are plans in place to use this to help define the geology majors program (contact: Sara), administering through all 4 courses and following a single cohort (over next 2 years).
   b) It could be interesting to connect with Christchurch on this – the program there (and the NZ post-secondary system in general) is very different from North America.
2) Climate Change (112, 340): largely validated, with minor work to be done refining questions.
   a) These are part of an extensive research project, and results are used in a variety of ways within the courses (contact: Sara).
3) Intro Lab (111): not officially validated, but a lot of work has already been done to refine these questions and it probably wouldn't take more than a small set of interview questions to finalize them.
   a) As course is changing hands and data haven't been utilized too heavily, further administration may not be useful but there is a large past data set (~5 years) which could be analyzed in new ways or summarized.
4) Computing in EOAS (211): not formally validated, but has been used since 2009.
   a) This is one of few cases where the instructors are still using their assessment even after the STLF is no longer here. It may be interesting to talk in more detail with Rich and Catherine about what they're getting out of it.

MEDIUM PRIORITY INVENTORIES

These are defined as medium priority mainly because they don't really have a champion, so to speak, but have potential (either in validation, previous work, or instructor interest). Or maybe they just don't necessarily have a plan to proceed.

1) Thesis-related projects (geologic time, landscape formation timescales (and corresponding confidence), spatial reasoning): most rigorously validated inventories but are under-utilized. It may be useful to think about why these were created initially.
    a) Geologic time – a concept that is integral to geology and is commonly difficult to grasp, with few dedicated assessments in the literature (and none at the upper level). Rhajiak, Jamil; Understanding Geological Time: A Proposed Assessment Mechanism for Beginner and Advanced Geology Students at UBC, 2009, https://circle.ubc.ca/handle/2429/6655
    b) Landscapes – a concept that is not covered in any published inventories (to our knowledge), yet was noted by instructors at UBC as one key aspect of geologic time that should be understood by students.
        i) When this is published (accepted for publication in May 2013, Journal of Geoscience Education), it would be a great idea to get this out into the department and in use. And if it doesn't get published, we'll probably need to get more data to make it publishable!
        ii) Jolley, Alison; Measuring Knowledge & Confidence of Landscape Identification & Formation, 2010, https://circle.ubc.ca/handle/2429/23321
    c) Spatial reasoning – a key skill that is needed for both structural and field geology, but is difficult and not obtained by students prior to studying these courses.
        i) Unfortunately, it was not used during boot camp in 2012.
        ii) Wong, Carrie; Performance on the Geologic Spatial Visualization Survey: A Comparison Between Junior and Senior Undergraduate Students, 2011, https://circle.ubc.ca/handle/2429/34237
2) Model Based Reasoning (212): used for a substantial research project on scientific thinking skills.
    a) FJ did run this in 2012 and is still (January 2013) considering how best to utilize results. For example, is it useful as a regular feature of the course, and if so why?
    b) One semester of data (2011-2) has not yet been processed, but it does exist.
3) Oceanography (372/373, 472): most extensively used for student feedback, but data should exist in raw form for all of them (pre-372 online diagnostic: see Susan or Maite, post-372/pre-373 review clicker questions: see Susan or Maite, pre/post 472 paper data: see Jason McAllister or Kristin).
    a) There is a lot of room to do some work with characterizing the oceanography suite by processing this data, if it were not difficult to track down. But is there a need or motivation for this? Is anyone interested?
4) Oliver Field School (328): previous data was collected almost solely for research purposes, however, with continued improvement of this course there is lots of room to grow.

Additionally, James Scoates has identified assessments as something that he would like to start back up in the upcoming semester for eosc331, eosc220 and eosc328. Thus far, this is just a passing thought that needs to be followed up with questions about intent and how best to make the data useful.

LOW PRIORITY INVENTORIES

These are defined as low priority mostly because they were never really utilized, or have been dropped with no indication of interest in picking them back up. Or, they would be very time consuming and probably aren't feasible (e.g., development of new assessments from scratch).

1) Saltspring Field School (223): in past, this was administered through MLB's own initiative. It was analyzed by her directly, so she saw and reacted to the results, as well as giving feedback to the students.
    a) Stopped doing so because she felt Carrie's thesis testing "interfered" with her own. This really isn't true, but with MLB being so busy anyway, this is probably not an effective place to target as there likely won't be much opportunity to respond to feedback.
2) Hydrology (329): only administered once (2010-2) and not validated.
    a) Not much use gotten out of it the first time around and no real interest for future (other changes are being made to the course that are unrelated to this).
3) Tectonics (332): not validated and although administered twice, the data were never really used.
    a) Brief feedback was given, but there wasn't anything substantial taken from the data.
4) Developing further instruments to attempt to get at habits and skills-type stuff would address the final piece of the puzzle to curriculum program objectives (e.g. more things like the MBR activity for 212).
    a) This could fit nicely into ENVR 200/300/400 and/or EOSC/ENVR 449.
5) It wouldn't be too hard to find validated (or semi-validated) questions for the 100-level classes (e.g. through GCI, SERC, or from Canterbury), but the administration would be more involved (large N) and would only be worth it if there was a defined outcome pursued.
6) These CIs focus heavily on geology, and to a lesser extent oceanography. Further CIs could be developed to address geophysics, atmospheric science, geological engineering, and environmental science.
    a) There is a strong need here but again it probably isn't feasible unless there is a real clear idea being pursued.

GENERAL RECOMMENDATIONS

1. Seeing how few instruments were used for course changes, it probably isn't realistic unless the instructor is highly committed and has a fair bit of time to work on it.
2. Exploring more ways to provide student feedback may be a useful avenue to pursue. Although this isn't the most typical use of CIs, it is a nice way to have the results help the students immediately – rather than only benefiting future students. It may make them keen to do their best and understand a small part of how educational research actually does help them.
3. Identifying areas of research more clearly and explicitly might be a good way to involve instructors if they are keen (helps with buy-in if they can see an outcome or product in academia).
4. Probably the best use of any RA/TA time would be to use past written responses to make multiple choice questions for as many of the inventories that are still in use (eliminates issues with time that data processing takes in the future). This just takes basic coding to find common distracters.
5. External collaboration is still of interest but coordination is a very challenging barrier for many reasons. We want to keep in mind the Christchurch NZ group, the GCI group, SERC and others. However, it is hard enough keeping our own local colleagues engaged let alone managing collaborations. Further progress will depend on someone to take the initiative and "drive" any collaborative efforts consistently through to some outcome.

TEMPLATES

A concept inventory processing spreadsheet (CIPS) for multiple choice items (partially manual but automates scoring process once correct answer key is entered).

- It has instructions for item analysis, gain calculations, and comparisons of all data and plots pre/post average scores by question on a bar graph.
- To incorporate SPESS data into the CIPS, there is a separate SPESS spreadsheet where that data first needs to be processed. It has instructions, and it uses some of Tom-Pierre's macros but none of the MATLAB or DOS operations (i.e., will work on any machine with just one spreadsheet). This is still a little time consuming, but manageable if the interest is there. Archived work can be found on EOS-SEI project backup hard-drives (the first was a 500Gb unit, filled to capacity as of Aug 10th 2011; the second 1Tb unit is still being used at Jan. 2013), and on the computers of Sara Harris or Francis Jones.

ADDITIONAL DOCUMENTATION

1) Individual assessment of each inventory, matched to course objectives, with general and specific recommendations (see "CI Individual Assessment" docx files in each respective folder)
   a) Needs followup by SH, FJ, BG; after looking at these notes and discussing (perhaps only briefly) priorities and directions.
   b) Was not done in great detail so I wouldn't make the suggested changes without pondering them a little bit, but can be used to help think about the quality of the questions being asked, and hence, the data that are produced
2) More detailed recommendation documents (begin with "Recommendations-")
3) Other brainstorming documents, that probably won't be helpful but were useful in me organizing my thoughts ("Objectives for NonSpecialist Courses" and "Strategies for Concept Test Incorporation").
4) (see "Summary of CIs" Excel file for a more detailed look at this).

OTHER WORK DONE

XX% of time was spent working on landscapes, and YY% on other projects. Other was: coding a couple of climate change questions for Sara (~2 days), making minor edits on Brett's 114 exams (2 hours), and connecting NZ work to here (5.5 hours; observing an exercise that I developed in Canterbury being run in 114 and reading reports and outcomes from the Canterbury projects).

CI Name	Course	07-2*	08-1	08-2	09-1	09-2	10-1	10-2	11-1	11-2	12-1	Contact	Res.	Ins.	Stu.	Con.	Valid.	Prior.	Use
EOS IntroLab	111											RM, SH, BG					2	1	C
Climate Intro	112											SH, EL					3	1	ALL
Computing EOAS	211											JC, RP, CJ					1	1	I
Model Based Reasoning	212											FJ, MJ					1	2	I, R
Mineralogy	220											RM, JS, JC, ML					3	1	ALL
Petrology	221											MK, BG					3	1	ALL
Saltspring (incl Spatial Reasoning)	223											ML, JC, SH					1	3	I
Ig Pet Diagnostic	321											MK, BG					3	1	ALL
Met Pet Diagnostic	322											GD, EL					3	1	ALL
Geologic Time - 326	326											FJ, SS					3	3	I
Oliver (incl Spatial Reasoning)	328											JC, JS					3	2	I, R
Hydrogeology	329											FJ, JC, RB					1	3	I
Landscapes IFT	N/A(330)											FJ, AJ					3	2	R
N.American Tectonics	332											BG, JM					1	3	C
Climate Upper	340											SH					3	1	ALL
Oceanography 1	372 (370)											SA, MM, EL					1	2	I
Oceanography 2	373 (371)											SA, MM, EL					1	2	I
Chemical Oceanography	472											KO, JMc, JC					1	2	I
Geologic Time - Jamil	N/A											FJ					3	2	R

	Pre-post not used/course not offered
	Processed
	Not processed

	Used for this at some stage
	Never used for this

Res. (Research):	Used by an STLF, student, and/or instructor to contribute to some sort of research project.
Ins. (Instructor Feedback):	Results looked at by one (or more) instructor(s) as general/specific feedback on learning.
Stu. (Student Feedback):	Results provided to students as general/specific feedback at relevant moments throughout semester, or CI itself used as activity/review.
Con. (Content/Curricula):	Results used by STLF and/or instructor to explicitly change course content or create new activity, lab, etc.
Valid. (Validation)^:	1-construct, 2-construct + content OR communication, 3-all 3 (one iteration), 4-all 3 (two iterations).
Prior. (Priority):	For further administration and/or data processing. 1-high, 2-medium, 3-low.
Use:	Recommendation for how it should be used (of 4 choices above).
*:	Indexed to calendar year (NOT "UBC calendar year") and semester number within that year.
^:	Validity definitions: 1) "construct" - experts have determined that topic measured is important; 2) "content" - expert has determined that questions measure what's intended (forced answer); 3)"communication" - questions are interpreted as intended by students. A&W (2011) suggest two iterations of student interviews (one with open ended, then test administration, then the second with forced answers).

Reference: Adams & Wieman, 2011, Development of assessments...IJSE, 33:1289-1312.

TABLE 1: SUMMARY OF ALL CONCEPT INVENTORIES IN EOAS (2007-2012) AND HOW THEY'VE BEEN USED.