Teaching, learning, and assessment in higher education: Using ICE to improve student learning

As published in the *Proceedings of the Improving Student Learning Symposium,* London, UK, *13*, 105-115. Imperial College, London, UK, September 2005. Oxford Centre for Staff and Learning Development.

Sue Fostaty Young
Queen's University
Kingston, Ontario
Canada

With increasing success, university professors are rising to the challenge of creating and implementing innovative teaching strategies that provide students with opportunities to engage in critical thinking, and creative problem solving as ways to enhance learning. Despite those innovations, students are still most likely to make choices about what and how to study based on how they will be assessed. Because students often regard grades as a "kind of currency indicating what teachers value", the influence of them swamps the effects of all other aspects of the curriculum (Boud, 1990, p. 103), including instruction (Wilcox, 1993, p. 7). Now, a pressing challenge for professors is to find and implement assessment methods that reflect and support valued learning processes as well as desired content outcomes. This paper outlines a model of qualitative learning assessment (ICE), with implications, and limitations for use in higher education and illustrates its utility through current-use examples.

Assessment has long been recognised as the single most influential factor in shaping what and how students in higher education choose to learn (eg Ames, 1992; Ramsden, 1992; Pintrich, Marx, & Boyle, 1993). In fact, it seems that no matter what innovative and engaging teaching methods are used, assessment will "swamp the effects of any other aspect of the curriculum" (Boud, 1990, p. 103) causing students to base their decisions regarding approaches to learning on how they will be graded, not on how they are taught. The influences of assessment on approaches to learning are so strong that Ramsden (1992, p. 187) insisted that assessment "defines the curriculum". A case in point serves to illustrate these contentions:

A group of in-service firefighters was required to complete a senior secondary level chemistry course as part of their ongoing professional upgrading. The experiential and inquirybased sessions were specifically designed to meet the group's professional needs and to reflect their day-to-day experience with chemistry. The firefighters were highly engaged and quite successful in their mastery of the material during the active-learning sessions and they actually looked forward to demonstrating their understanding of the course material through the mid-term exam.

After being faced with a 50-item multiple-choice exam that was based entirely on the text used as a reference in the course, the group presented a united front in their refusal to participate in any further activity-based sessions. They acknowledged that the learning gained through past sessions was superior to that likely to be gained from more traditional, lecture-based classes, but argued that to do well on the exams, the approach to the material needed to be changed. They were willing to give first priority to the pursuit of grades rather than to the pursuit of learning, something that Wilcox (1993) acknowledged as a major obstacle to deep learning in higher education. Rather than ask that the method of assessment be brought into alignment with the teaching methods that enabled deeper learning, the group asked for the teaching methods to be brought into alignment with the assessment. The firefighters evidently perceived that the method of assessment, and not the teaching, set the criteria for course expectations and did indeed define the curriculum for them.

Brookfield (1995) cautioned teachers to consider that what we choose as evaluative criteria serves as a public declaration of what it is that we value and stand for in our work. Certainly, the firefighters seemed to interpret assessment as such. Boud (1990, p. 103) went further in stating similar beliefs in very practical terms, asserting that students view "grades as a kind of currency indicating what teachers value".

Evidence to support those contentions emerges, without fail, on the first day of every course I teach. When the course syllabus is distributed, students will, almost without exception, turn immediately to the last page of the document to review the assessment plan for the course – to determine, as Gass (2003) would explain, what they would "get paid for" in the course. The information on that last page helps students to determine their focus and responsibilities and to make determinations about approaches to course material. After all, preparing for multiple-choice exams requires an entirely different approach from preparing a seminar presentation, which in turn is entirely different from approaches to material for laboratory work or practica.

Aside from having the potential to reveal what it is that we value, our assessment plans can also serve to give cues about our beliefs about the nature of learning (Schrag, 1992). I wanted to ensure that assessment in the courses I teach conveys a value for personal meaning-making and for deep approaches to learning. Moreover, it was important that assessment methods were congruent with more of a cognitive transformative view of learning than a behavioural one. I wanted to find an effective and efficient way to help explain to students why more of the same is not always better, or more accurately, that when professors ask for "more of an answer" we are not looking for more words or more pages but rather, looking for more depth and greater expertise.

Taxonomies of learning can help inform decisions about teaching, learning and assessment and aid in sharing views and expectations about learning, but finding one that meets faculty's expectations for ease of use can be challenging. Some taxonomies, such as Bloom's (1956), compartmentalise learning into separate domains as if cognitive, psychomotor and affective learning are discrete processes, mutually exclusive from one another. Having separate taxonomies for each of the different domains of learning, and each with at least 5 levels of learning, makes the model unwieldy and reduces its portability for effective use in the classroom.

Biggs' and Collis' (1982) SOLO taxonomy follows a Piagetian view of cognitive development, and depicts learning as growth from novice toward expert, something that resonates with many faculty members. Learning is viewed, not as the accumulation of discrete information but rather, as a transformative process where the learning as well as the learner is altered through development. While the taxonomy is useful in developing educators' and researchers' understanding of how learning occurs, the five levels of learning (prestructural; unistructural; multistructural; relational; and extended abstract) make it difficult to use in anything less than a formalized way. Faculty members report that while the model does reflect the way in which they believe learning occurs, students with whom they share the model spend more time trying to decipher the model than they do on their own learning of the course material.

The ICE Approach

In an effort to condense the cognitive transformative literature into a portable model that would be accessible and easy for both teachers and students to use, Wilson (1996) introduced the ICE approach. The approach is portable in that it is easily remembered and called to mind and it is applicable to learning across domains and levels of education. It is accessible in that the language of the model makes it easy to share and understand, so much so that Grades 4 and 5 students were recently responsible for explaining the model to a group of pre-service teachers enrolled in an undergraduate assessment course. The model is simple without being simplistic.

Elaborating upon Wilson's introduction, Fostaty Young and Wilson (2000) presented *Ideas, Connections,* and *Extensions* (ICE) to represent three different levels of learning growth from novice through to competence and expertise. *Ideas* represent the building blocks of learning. They are the fundamental, discrete pieces of information that make up the basics of new learning. Some teachers describe *Ideas* as being only information, something students acquire then posses. They include facts, definitions, vocabulary, steps in a process and discrete skills. Any reiteration, or recall of information from a textbook, notes or lecture can be said to be a demonstration of *Ideas* level learning.

Connections are of two kinds: those made at the content level and those that may be said to be personal meaning-making. *Connections* at the content level are demonstrated when students are able to articulate relationships among discrete *Ideas.* When students are able to describe cause-and-effect relationships, articulate the relationship between or among concepts, or when they are able to successfully blend two or more discrete skills into a fluid, efficient movement, they are demonstrating *Connections* at the content level. *Connections* at the more personal, meaning-making level are demonstrated when students are able to relate their new learning to what they already know. It is during this phase of personal meaning-making that learning appears to take on a new dimension in that it seems to become more easily retrievable and longer-term than learning at the *Ideas* level.

At the *Extensions* stage, new learning is created from old so that students are able to use it in novel and creative ways that may well be quite far removed from the original learning context. The learning becomes internalized to such a degree that it helps students answer extrapolative questions, articulate implications, and anticipate outcomes. *Extensions* are referred to by some as the AHA! phase of learning and by others as the “so what?” phase. The “so what” question is the one that is often left unasked: So, now that you know what you know, what difference does it make to the way you see the world and to what you can do? Students reaching *Extensions* are able to answer those questions.

The questions we ask have an interesting effect on the answers we receive. To a great extent, the questions asked determine the answers supplied. Closed questions inviting reiterations from texts and multiple-choice questions cannot afford students the opportunity to demonstrate *Extensions* even if they are capable of making them. Invitations by way of questions need to be made accessible to students at each level of learning development as well as being *extend*able by those who are able (Fostaty Young & Wilson, 2000). That is not to say that every assessment event must include opportunities for demonstrations at every level of learning. There are times when it is important to determine that students have acquired the basic, fundamental aspects of the material before moving on. In such cases testing entirely at the *Ideas* and content-*Connections* levels may be exactly what is warranted. It is important, however, that we become conscientiously aware of the types of opportunities we are providing and if we are providing sufficient opportunities for the learning we hope to elicit. As one Engineering professor remarked, “What is measured is what is produced.” Similarly, what is asked is what is answered.

Having a systematic framework that facilitates our ability to plan and review exams increases the likelihood that we are actually asking the questions we would like students to be able to answer. Tables of specification, coupled with using ICE as a framework, function as effective tools for such tasks. The tables are a type of spreadsheet that help clarify the elements of a course that are to be highlighted on a particular exam. Typically, they are in the form of a grid with course-content areas along one axis and levels of learning along the other. Topics for the test are decided upon and listed, a total point value for the test determined, and then a decision is made as to the relative weighting of each topic. Finally, decisions are made as to the level of learning that each topic area will target. The following example from a course on contemporary issues in business serves to illustrate:

Content Area	***Ideas***	***Connections***	***Extensions***	**Totals**
Privatization		10	10	**20**
Mergers and Acquisitions	10	10		**20**
Gender Discrimination	20	10		**30**
Ethics			30	**30**
Totals	**30**	**30**	**40**	**100**

Figure 1

The table indicates that the exam is relatively evenly balanced across the listed content areas but that *Gender Discrimination* will deal primarily with fundamental details whereas it is likely that students will be expected to be able to interpret, postulate and problem solve in the section on *Ethics.*

Interesting to note is that even if the totals for each content area remain the same, by changing the grade-weights among the *Ideas, Connections* and *Extensions* cells, the difficulty level of the exam will be significantly affected. To ensure that exams that are made available to you from outside sources are meeting your expectations as well as the learning needs of your students, you may choose to use a table of specifications to conduct an analysis. Review each question on the existing exam and categorize it as one that invites responses at *Ideas, Connections* or *Extensions* and make note of the grade values associated with each, recording your findings on a blank table. You may then be better able to determine the appropriateness of the test for its intended purpose. Additionally, you may be able to systematically cull the questions that were difficult for the wrong reasons - not because they required more expert levels of learning but because they were poorly posed.

ICE and Grading

While ICE was originally designed and developed as a formative assessment tool to help teachers and students plan and improve learning, it is increasingly being adapted for use as a summative tool. Faculty, in particular, seem to view *Ideas, Connections,* and *Extensions* as representations of *Acceptable, Better,* and *Ideal*, levels of learning growth, respectively. They frequently understand *Ideas* to be the fundamental elements of the course content that students must be able to demonstrate if they hope to be at least minimally successful in the course. Given that most faculty members also express the opinion that acquisition of subject content is not enough and that their goals for their students include creative problem-solving and novel use of knowledge, they are unwilling to assign much more than barely passing grades to demonstrations of learning at the *Ideas* level. Because learning at the *Connections* level is deeper, more meaningful learning, it is more highly valued and faculty members consistently award demonstrations at that level higher grades. Since demonstrations at the level of *Extensions* are almost unequivocally regarded as ideal learning in higher education, many faculty members reserve the highest grades for demonstrations at that level.

Sharing ICE with Students

As the result of a recent study, Rust, Price and O'Donovan (2003) concluded that student learning can be improved significantly through developing their understanding of assessment criteria and processes through structured intervention. Their intervention included providing students with a criterion-referenced grading grid, or rubric, that included a matrix of detailed performance indicators across grade levels. They also provided an opportunity for students to attend a one-of workshop to familiarise them with expectations for the course's major assignment. Findings suggest that students who attended the workshop not only did better in the course than their non-participating cohorts but that the positive effects were long term and seemed to transfer to other, similar learning situations.

Through a much less structured but apparently equally successful approach, faculty members who have become familiar with ICE are sharing the model with their own students, often as early as the first day of class. Typically, the process begins with a brief description of *Ideas, Connections,* and *Extensions* and illustrative examples from their own disciplines or from the wider world. Such examples serve to practically illustrate the model in a way that makes sense to the students. One example that has been widely borrowed (Fostaty Young and Wilson, 2000, p.54) was supplied by a Business professor's use of analogy to convey his conceptions of "adequate", "good" and "exceptional" approaches to case study analysis. In discussing the variability of his students' completed analyses, the professor likened the case study to a broken toaster. Some students responded to the task by itemising all the elements of the toaster (case study) that were malfunctioning. He conceded that while students had done a good job at doing that, some had stopped there, making no attempt to go further (*Ideas).* Another group of students, he said, also indicated which parts of the toaster/case-study were malfunctioning but they then also proceeded to postulate how those parts were affecting the mechanisms that were still operational (*Connections*). Further, a smaller proportion of the class pointed out the parts that were broken, indicated what effects those were having on the functioning parts and then, suggested how the toaster might be repaired and how future malfunctions might be prevented (*Extensions*).

Another innovation that some faculty members are using to develop their students' awareness of differences in the quality of cognitive output is in inviting students to read with a critical eye. As a preliminary exercise students are assigned an article and asked to use different coloured highlighters to identify the *Ideas, Connections,* and *Extensions* expressed by the author. In addition to helping students learn the differences among the levels of the ICE model the exercise also helps to reinforce the use of a common classroom vocabulary to discuss learning and assessment criteria.

Sharing assessment criteria with students, and in some cases, developing the criteria *with* students is a critical part of the ICE approach. Most faculty who are using the approach have developed rubrics as a means of sharing those assessment criteria with their students. The rubrics are typically grids with no more than 5 content elements listed in a column on the left and the levels of ICE in rows across the top. In each corresponding cell, qualitative descriptors are provided that outline expectations at each level. The following is a rubric designed to assist students in their approach to an assigned discussion paper in a geology course (Remenda, 2005).

Element	***Ideas***	***Connections***	***Extensions***
Content	- Identifies the hypothesis - Statements are accurate in terms of the assigned paper - The summary provided accurately depicts the content of the paper - Provides accurate definitions for the vocabulary used in the assigned paper	- Draws attention to the relationship among topics within the paper - Provides real-life examples to illustrate concepts - Connects the content of the paper to course topics; explains the applicability of the data - Considers the suitability of the research methods - - Considers how well the data supports the conclusions	- Extrapolates from the paper to other situations - Evaluates the applicability of the hypothesis and why it is of general interest - Relates topics from the course and the paper to other disciplines or to the field

Figure 2

Note the characteristics of the rubric. Learning is described only in positive terms and in terms of what is in evidence, *not* in terms of what is preventing the piece of work from being judged at a higher level. The learning demonstrations listed under *Ideas* are the professor's basic, minimum requirements for success on this assignment. The descriptors under *Connections* and *Extensions* represent qualitative rather than quantitative criteria through which the work will be assessed. The rubric conveys a view of learning as a qualitative, cognitive transformative process rather than as a quantitative, behavioural one. It enables us to move beyond learning that is content-based and beyond a reliance on assessment that focuses on knowledge, recall and demonstration. It instead turns our attention to the interaction between the learner and the content and places emphasis on learning as personal meaning-making.

During the deeply reflective process of articulating levels of learning outcomes, many faculty members become acutely aware of the implications of using ICE as a framework for assessment. One such implication involves the necessary and concomitant shift in ongoing teacher/student interactions. Instructional methods necessarily evolve from presentation of content material to creating opportunities for meaning-making and students' active engagement with the course material.

Creating Rubrics to Increase Understanding of Learning Expectations

Creating qualitative rubrics is much more difficult and time consuming than creating the more typical quantitative rubrics that outline criteria of "few", "some", "many" and "almost all" or "inconsistently", "generally" and "consistently". It requires a focused attempt to specifically articulate what it is that we seem to intuitively and tacitly *know* when reviewing students' work and there are a variety of ways in which to approach the task.

One method is to review a set of work that you have already graded. Select a few to which you awarded barely passing grades, others to which you awarded mid-range grades, and some to which you awarded high grades. Next, set out to outline characteristics of each group that led you to award the grades you did. Confine the descriptors to statements of evidence; try not to describe the work in terms of what it is missing. It is likely that you will find patterns among the groups of papers or assignments that will enable you to create an explicit rubric representative of the tacit one you had been using.

If you do not currently have access to an already graded assignment-set, you may use another equally effective approach to construct a grading rubric. Select a specific assignment that you are prepared to set for your students and list what you consider to be the minimum requirements for success. That is, what are the basic criteria that students absolutely must fulfill in order to achieve a passing grade? After listing all the minimum requirements for success, group them into similar categories, then determine what each category should be called. Continue filling out the remaining cells using descriptors that represent differences in what students are able to accomplish with the basic *Ideas* rather than differences in the number of times they demonstrate the same skills. For example, in a rubric for mathematical problem solving, rather than using descriptors such as *demonstrates limited understanding*; *general understanding*; and *overall understanding* for each of the levels of learning, consider descriptors that detail the desired processes and approaches to learning such as: *makes valid observations based on available information*; *articulates the relationship between the known facts and the problem to be solved*; and *is able to explain his/her thinking using ideas from different situations to help explain new ones.* The qualitative descriptors provide a blueprint for learning for the students without confining them to specific content or to behavioural demonstrations.

A young Biology professor preferred to use an altogether different approach to rubric construction. She decided to involve her 4th-year students in creating the rubric for their seminar presentations. She invited the 16 students in the group to list what they considered to be the basics of a good presentation – what they would consider to be a passable seminar. As elements were proposed and agreed upon she began to insert them onto a grid on the chalkboard. Next, they generated the qualitative descriptors for a presentation that they felt would be worth a B-grade then continued on to describe a stellar performance, one they all felt would be worthy of an A. There was much discussion but through a process of general consensus, the group developed a grading rubric that included descriptors at three levels (*Ideas, Connections* and *Extensions*) for elements such as seminar content; context; use of media; and facilitating group discussion. The professor typed up the rubric and solicited students' feedback before declaring it ready for use. Using the rubric they had generated as a group, they all, students and professor, assessed each seminar.

While the generation of the rubric in the Biology seminar course was time consuming, both the professor and the students now consider it a worthwhile investment. Not only did the students claim that the process had helped them tremendously in the seminar course by clarifying expectations, but they reported a transfer effect and felt they were able to apply the principles of ICE to their learning in other seminar-style courses. Most notably, other faculty members within the department have commented that students from the seminar course are beginning to outperform their counterparts in the rest of the program.

In sharing ICE with students professors report a developing consciousness and greater awareness in themselves and their students about learning, about how to foster it, and how to demonstrate it. Students become better able to self-assess and, as a result, can better plan for improvement.

Using ICE Across the Disciplines

ICE's portability certainly seems to increase its utility. The approach is being successfully used across disparate disciplines for skills training and other activities typically referred to as psychomotor learning as well as for cognitive learning. Departments of physical and health education, rehabilitation therapy, nursing, languages, history, fine art, and applied sciences are all finding uses for ICE within their disciplines. Many are discovering that ICE provides the necessary vocabulary and supporting framework for offering feedback and guidance, especially for difficult-to-assess areas of study and for decision-making purposes:

- Rather than use a behavioural checklist to assess occupational therapists' work with clients, the department has devised an ICE rubric that takes into account and values the critical analysis and expertise than can result in abandoning rule-sets in favour of adhering to maxims to guide action.
- Students in fine arts courses are learning the specifics of the previously tacit criteria their professors were using to assess their artwork. Together, students and their professors are reviewing pieces of work in specific genres and describing the qualities of brush technique, shading, mixing, and balance that distinguish exceptional art from amateur work and then creating rubrics with which to plan, and then assess, course work.
- Recently a professional school adopted ICE as the framework within which to evaluate letters of application from prospective students. In the past, admission

> committee members had difficulty distinguishing among the almost uniformly stellar students. Using ICE, they are able to distinguish the letters that itemise evidence of academic suitability from those that demonstrate the writer's awareness of the potential impact of that academic excellence on future success and suitability to the field.

The ICE approach seems to resonate with faculty and students alike. It is easy to use and seems to be applicable to almost any learning activity. Because it is not content-specific, professors are discovering that it is not necessary that every student produce identical products, that evidence of learning can be gleaned in a variety of ways.

Reaction to the approach has been overwhelmingly, though not unanimously, positive. Nonetheless, even those who find no reason to adopt the ICE approach outright find that the model's accessibility furnishes the vocabulary that facilitates discussions about assessment and the aims and values that inform it. In addition, the relative simplicity of the model allows for easy modifications and adaptations to be made that might enable a more accurate representation of individual professor's values and expectations for learning. The model's simplicity, and the language it provides, help to facilitate discussion not only about assessment in higher education but about teaching and learning as well.

References

Biggs, J. B., & Collis, K. (1982). *Evaluating the quality of Learning: the SOLO Taxonomy.* New York, NY: Academic Press.

Bloom, B.S. (9156). *Taxonomy of educational objectives: The classification of educational goals.* New York, NY: David McKay.

Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education, 15*(1), 101-111.

Brookfield, S. D. (1995). *Becoming a critically reflective teacher.* San Francisco, CA: Jossey-Bass.

Fostaty Young, S., & Wilson, R.J. (2000). *Assessment and Learning: the ICE approach.* Winnipeg, MB. Portage and Main Press.

Gass, L. (2004). Invited lecture, Queen's University, Kingston, Ontario, Canada

Ramsden, P. (1992). *Learning to teach in higher education.* London and New York: Routledge.

Remenda, V. (2005); Queen's University; Geological Sciences; personal communication.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment and Evaluation in Higher Education, 28*(2), 147-164.

Pintrich, P.R., Marx, R.W., & Boyle, R.A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*, 167-199.

Schrag, F. (1992). Conceptions of knowledge. In: P. Jackson (Ed.). *Handbook of research on curriculum* (pp. 268-301). NY: MacMillan.

Wilcox, S. (1993). Evaluating learning in higher education: A guide to the literature. *Higher Education Group Annual*, 2-22.

Wilson, R.J. (1996). *Assessing students in classrooms and schools.* Scarborough, ON: Allyn and Bacon.