University of Waterloo Teaching and Learning Conference, April 30, 2015

Making Your Students' Learning Visible: *How Can We Know What They Know?*

Linda B. Nilson, Ph.D., Director
Office of Teaching Effectiveness and Innovation
Clemson University, 448 Brackett Hall, 321 Calhoun Drive
Clemson, SC 29634 USA * 864.656.4542
nilson@clemson.edu * www.clemson.edu/OTEI
www.linkedin.com/in/lindabnilson

Participant Outcomes

- To give good reasons for measuring learning.
- To explain course alignment and its essential role in student learning
- To write good, assessable student learning outcomes and assess your students' achievement of them

- To assess most higher-level thinking skills with objective items.
- To design/adapt assignments, tests, and survey instruments to measure course-level student learning in several ways.
- To evaluate the merits of these measures.
- To express learning in simple numerical terms (to facilitate use).

Why Measure Learning?

- To assess a new method or approach
- To add evidence of your teaching effectiveness in your reviews
- To help depts gather data for program review/accreditation
- Just to know …..

Aligned Course (and Curriculum) Design

Appropriate Assessments of Students' Achievement of the Outcomes

(the measurement of students' progress toward the ends)

↑

Learning Opportunities/Teaching Methods to Help Students Achieve the Outcomes

(the means to the ends; the "tools for the job")

↑

Student Learning Outcomes

(the foundation, the ends of instruction)

What Are Good Student Learning Outcomes?

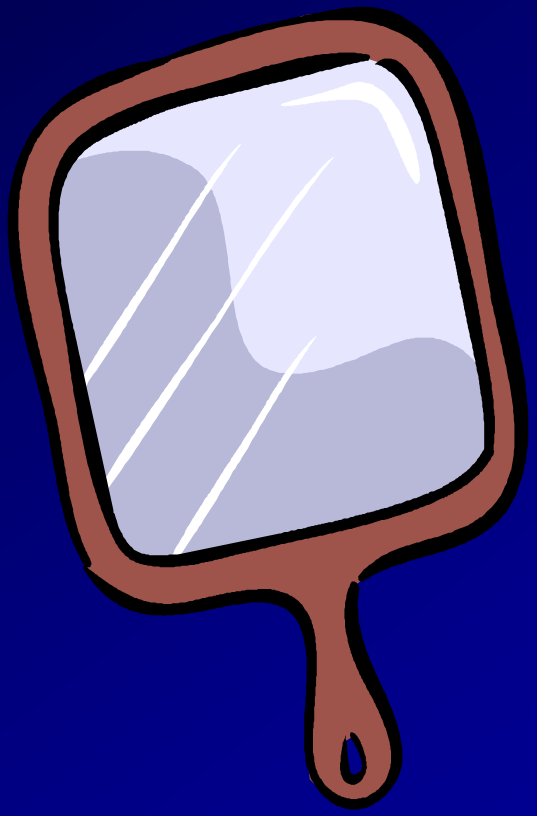
- Statements of what your students should ***be able to do*** by end of unit or course.
- "Performances" that you can ***observe*** and ***set standards for*** so you can ***assess*** them—***active verbs***
- **Not** internal states of mind: "know," "feel," "learn," "understand," "appreciate"

General Categories of Learning Outcomes

- Psychomotor
- Affective
- Social
- Ethical
- Cognitive *

Assessments Should *Mirror* Outcomes.

Outcome



Assessment

Golden Rules of Assessment

1. If you want your students to be able to do X, Y, and Z, have them do X, Y, and Z to assess whether they can.

2. Before assessing summatively, assess formatively to:

- Give students **practice** with feedback from you, their peers, or computer program. Practice = in-class activities and assignments
- Get frequent **feedback** for **yourself** on their progress.

Summative Assessment Instruments

- Objective test items

  = fill-in-the-blank (completion), true-false, matching, multiple choice, multiple true-false

- Student-constructed work

  = writing assignments, essay test questions, oral presentations, multimedia programs, projects, research reports, designs, artistic works, portfolios

Multiple True-False Items

- Superior flexible, efficiency, reliability
- Easier & quicker to develop
- More challenge, no process-of-elimination
- Stem must be clear.

MC and MT/F Can Assess Higher-Order Thinking

Interpretive exercises (aka *scenario, context-dependent, simulation-like* MC or MT/F items)

= **series** of MC or MT/F items based on a new*, realistic stimulus (experiment, statement, passage, mini-case, graphic, etc.) that students must read/examine to answer the items correctly

** New to the students*

Interpretive Exercises Can Assess:

- Interpretation
- Generalization
- Inference
- Problem solving
- Conclusion drawing
- Comprehension
- Application
- Analysis
- Synthesis
- Evaluation

… But Not

- Abilities to communicate, create, organize, define problems, or conduct research. ***Only student-constructed work can assess these.***

Guidelines for Writing Interpretive Exercises

- New stimulus, but students need **prior practice** in the thinking skills assessed
- Few interlocking items
- Length/complexity of stimulus ≈ # MC or MT/F items possible
- Be creative with stimulus!
  chart, graph, map, picture, diagram, drawing…

Measures of Learning

	Indirect	Direct
End-of-course-only	Perceived Student Learning Gains Instruments (self-regulated learning activity)	Integrative Essay or Journal Entry (capstone paper or final exam) Targeted Essay (capstone paper or final exam)
Pre-and-Post-Test	Knowledge Surveys (self-regulated learning activity)	1st-Week Writing (ungraded) and Correction Exercise (final exam) 1st-Week Essays (ungraded) and "Value-Added" Essay Final Exam 1st-Week Final Exam (ungraded) and Final Exam

- End-of-course-only – indirect
  - Perceived student learning gains instruments
- End-of-course-only – direct
  - Integrative essay or journal entry
  - Targeted essay
- Pre- and post-test – indirect
  - Knowledge surveys
- Pre- and post-test – direct (pre-test ungraded)
  - 1st-week writing + correction exercise final exam
  - 1st-week essays + “value-added” essay final
  - 1st-week final exam + final exam

End-of-Course-Only Indirect

Perceived Learning Gains

- Student Assessment of Learning Gains (SALG) survey instrument www.salgsite.org
  - Items on *most effective course elements* (class activities, assessments, learning methods, labs, resources) and *perceived learning gains* (general, understanding concepts, acquiring skills, developing positive attitudes, integrating info)

- Validated: In testing, r = .41 between students' SALG scores and their scores on final exam.
  But in specific topic areas r = .0 - .49 between students' SALG scores and their scores on corresponding subsections of final exam.

To represent learning numerically:

Average score across relevant perceived learning gains items

Strengths?

Weaknesses?

Weakness:

- Recent evidence that students don't always perceive their learning accurately (Bowman, 2011; Porter, 2012; Weinberg, Fleisher, & Hashimoto, 2007; Weinberg, Hashimoto, & Fleisher, 2009)
- Students rate their skills and work higher than faculty do, especially non-science and intro-level students. (Falchikov & Boud, 1989 – meta-analysis)

End-of-Course-Only Direct

Integrative Essay or Journal Entry

- Students review course material and draw their own conclusions about it, its value, and their learning. (Atlas, 2007 in expository writing course)

Targeted Essay

- For example: Answer to job interview question for dream job: “What are the most important things you learned in your XXX course? Demonstrate your skills in applying XXX.” (Weimer, 2007 in comm studies course)

To represent learning numerically:

Average numerical score (grade) of integrative or targeted essays

Strengths?

Weaknesses?

Weaknesses

- Instructor can stack grading results.
  - Peer reviewers should see the question direction, grading rubric, and sample essays.
- Impossible to identify and remove course-related knowledge that students had coming into course.

Better questions to suggest?

Pre- and Post-Test Indirect

Knowledge Surveys

- Series of questions or tasks covering the material – *knowledge and skills* – of an entire course or unit (from outcomes, exercises, old exams, etc.)
- Different levels of thinking
- Answer = *students' perceived ability* to answer question or perform task

Knowledge Surveys
Examples of Answers

a) I do not understand the question, I am not familiar with the terminology, ***or*** I doubt that I can answer the question well enough to earn a passing grade.

b) I understand the question and 1) I think I can answer at least half of it correctly, ***or*** 2) I know where I can find the correct answer within 30 secs.

c) I am confident that I can answer the question well enough to earn a passing but no higher grade.

c) I am confident that I can answer the question well enough to earn a high grade.

Strengths?

Weaknesses?

Weakness: Students don't always know what they do & don't know.

- They overestimate their abilities (except the best students) when they know the least.

Compression effect on learning measure

Less in engineering, sciences, and health fields, which use terminology students know they don't understand.

To represent learning numerically:

- Average difference in pre- & post-test confidence ratings across items

- Calculation of gain in confidence betw pre- & post-test survey (equations later)

Pre- and Post-Test Direct

Scientifically Strongest

1st-Week Writing + Corrections

- *1st week*: Ungraded writing assignment on key concepts, principles, techniques, processes, causes, effects, etc.
- *Final*: Letter to "pre-class self" correcting errors, poor reasoning, misconceptions, etc. in 1st-week assignment.

(Griffiths, 2010 in criminology course)

To represent learning numerically:

- Average numerical score (grade) of students' corrections final.

- Score the 1st-week writing as well as corrections final and calculate learning gain between pre- and post-test assignments (equations later).

1st-Week Essays + "Value-Added" Essay Final

- *1st week:* Ungraded take-a-stand essays in class (answers will be brief)
- *Final:* Rewrite essays with supporting evidence and critique of 1st essay, even if no change of position.
  - Developed & used by Dr. John ("Mike") Coggeshall, Professor of Anthropology, Clemson University

To represent learning numerically:

- Coggeshall grades final essays on amount of supporting evidence. He reports % of class that develops and backs argument w/evidence (his major ultimate learning outcome).
- Score 1st-week essays as well as final essays and calculate learning gain between pre- & post-test essays (equations later).

1st-Week Final Exam (ungraded) + Final Exam

Give students the final exam *twice*:

- *1st week:* Ungraded in class (they won't need very long); only score them
  - Don't tell students this is the final until afterwards. Harm done? They may recall what is important for them to look for and learn. *Not* for online courses.
- *Final:* as usual for grades

Strengths?

Weaknesses?

To represent learning numerically:

Debate on calculating learning gain

(Both measures eliminate students' pre-course skills & knowledge.)

$$\frac{(\text{Post-test\%} - \text{Pre-test\%})}{\text{Pre-test\%}} \times 100$$

By what percentage did students increase their knowledge of the course subject matter during the course?

vs.

$$\frac{(\text{Post-test\%} - \text{Pre-test\%})}{(100\% - \text{Pre-test\%})} \times 100$$

$$= \frac{\textit{Actual gain}}{\textit{Possible gain}}$$

= "average normalized gain"

What percentage of all the (tested) course subject matter did the students actually learn during the course?

Example of Interpretive Difference

$$\frac{75\%\ \text{(final aver)} - 20\%\ (1^{st}\text{-time aver)}}{20\%\ (1^{st}\text{-time aver)}} = \frac{55}{20} = 2.75$$

x 100: "Students increased their knowledge of the subject matter by 275% during the course."

vs.

$$\frac{(75\% - 20\%)}{(100\% - 20\%)} = \frac{55}{80} = .6875 \quad \text{x } 100$$

"Students learned 68.75% of the knowledge & skills that they could have in the course."

Which measure(s) is(are) your preference(s) for your reviews, and why?