

Classroom Assessment vs. Student Satisfaction

This is an article much in agreement with, but taking a different angle from, a recent Opinion piece, “Evaluation of our courses”, by Steven Zucker, which appeared in the August 2010 issue of the *Notices* (page 821).

Zucker raises an issue I have cared greatly about: our so-called course evaluations are not evaluations at all. Administrators interpret (and often use) the scores as measuring instructor effectiveness. Yet such scores cannot reflect how much students have learned. The scores encourage us to become popular teachers rather than good educators.

Testing course effectiveness requires deeper ideas than asking students how they liked the course near its end. No matter how you phrase the questions or what you ask, such a survey cannot provide reliable information about course effectiveness. By an effective course I mean one that enables students to use the material later on.

Fifteen years ago, I got extraordinarily high survey scores. Yet my vector calculus students did very poorly on final exams. In 1995, backed by a Sloan Foundation grant, I developed an Internet-based student assessment tool: “Interactive Questionnaires” (IQs). An IQ is an enhanced interactive quiz.

Here’s how it worked. I sent IQs to students in email. By applying one instruction, a student could open the IQ as a UNIX program. It would guide the student on problems he or she likely couldn’t have done independently. When the student was satisfied, the program automatically mailed me the student’s responses. My system produced formatted “interaction portfolios” from these responses. I could evaluate them electronically to assess the entire class on a finely parsed topic or an individual’s progress in the steps to topic mastery.

I could do this without losing class time, long before final exam failure. From this I found when and why my class’s modest early mastery could mysteriously disappear—rather than grow—as the final approached.

My published paper, at the URL <http://www.math.uci.edu/~mfried/edlist-tech/gold02-08-98.pdf>, gives example questions. It shows how I used interaction portfolios to produce IQ reports. You can see how the email system—“Problem of the day”—coordinated daily classes with IQ intercessions. IQs got students to write mathematics in a controlled setting. For IQ pieces that stretched students beyond a one-word answer, I could electronically batch-grade: an activity with no resemblance to the onerous task of inspecting cryptic scribbles on paper.

I designed IQs to engender step-by-step thinking. I added nuance to what most UC administrators knew of a particular minority student (see <http://www.math.uci.edu/~mfried/edlist-tech/gold02-08-98.html>). My story informed administrators that he was the only

exception to vector calculus totally wiping out minority students from participating in mathematics, science, and engineering.

I have visited prestigious universities. Their classes included some students whose mental energies allowed them to digest and analyze at impressive rates. Such students may quickly imitate the instructor’s “analyzing” with less training than I typically saw at UC Irvine. Zucker apparently sees that phenomenon at Johns Hopkins.

Responding to classes at different levels is part of the teacher skill. IQ questions must change according to class level, especially to take advantage of the additional skills students acquire using them. Yet, one change baffled me.

Despite documentable student progress on critical topics, my survey scores dropped from 6.2–6.7 to 3.5–4.5 (on a scale of 1 to 7). That shook me at first. (I am told that carefully controlled studies have shown negative correlation of survey scores to learning.) My demeanor and classroom presentation, even my graphic illustrations, did not change from pre- to post-IQs. Yet, post-IQ students saw me as less sympathetic to their trials with a difficult course.

Compatible with Zucker’s experience, the IQ reports led students to see they could work harder. Many of my students (certainly not all) interpreted that as a negative. Students at elite institutions might feel that using a tool like IQs is not necessary—a little like spoon-feeding.

Less necessary maybe, but it isn’t like spoon-feeding. As I became confident in the response to IQs, I saw students develop confidence that classes made sense, something many told me they doubted before this experience.

Having instructors at other institutions use IQs would allow comparing how the tool works for students with differing initial aptitudes. Teachers would get different results, but the improvement would be in student accomplishment, not in instructor likability.

My tools should go through a Web interface with security controls. Higher administration money backing such a development would take the onus off each instructor. I see that changing how students respond to working harder.

IQs provide, as could other assessments, a rethinking of how we document what our classes accomplish. As an alternative to polling students at the end of the semester, they can measure the ongoing effectiveness of a long course.

—Mike Fried
Emeritus Professor
University of California
at Irvine
mfried@math.uci.edu