Clarification to Points in “Correlating Student Knowledge and Confidence Using a Graded Knowledge Survey to Assess Student Learning in a General Microbiology Classroom”

To the Editor,

Favazzo, Willford, and Watson (2) merit congratulations for their paper and for incorporating knowledge surveys into an ongoing assessment program. After making good use of our early paper (3), they constructively extended it by working to deduce the relationship between self-assessed competency and demonstrable competency.

In practice, measuring self-assessment accuracy is not simple. My recent work with colleagues has increased understanding of some challenges that are useful to share here.

1. Data generated from both instruments used to obtain the paired measures must have a confirmed reliability to permit meaningful comparisons.
2. Reliability depends upon both the quality of the instruments and acquiring a sufficiently large database.
3. Both instruments must communicate the same construct so that participants clearly understand the actual challenge that they are expressing their self-assessed abilities to meet.
4. Competence and confidence have no established units, so the current default measures are scores reported in percentages. The measurements constitute arrays bounded by 0% and 100%, which adds numerical complications.
5. The capacity to overestimate or underestimate self-assessed competency is not uniform across all participants.
6. The effects of these challenges on the graphics commonly used to argue for conclusions are not obvious.

My current co-workers and I now consider both test and knowledge survey data as consisting of the measure sought (signal) and noise (mostly random variations). When noise is random, and the signal of self-assessed competency is systematic, averaging paired measures item-by-item should attenuate noise and allow better definition of the signal. Thus, collective item-by-item means rather than student-by-student correlations may more clearly show the relationship between self-assessed competency and actual competency.

Figure 1 from our ongoing work (http://serc.carleton.edu/earth_rendezvous/2015/program_table/abstracts/101047.html) employs self-assessed and demonstrated competency and knowledge survey data as consisting of the measure sought (signal) and noise (mostly random variations). When noise is random, and the signal of self-assessed competency is systematic, averaging paired measures item-by-item should attenuate noise and allow better definition of the signal. Thus, collective item-by-item means rather than student-by-student correlations may more clearly show the relationship between self-assessed competency and actual competency.

One statement by Favazzo et al. regarding our work merits clarification:

“…when students do not attempt to answer questions, they have significantly greater confidence in their ability to answer those questions. Thus, when measured alone, confidence is inflated. This finding may help to reconcile the disagreement between Nuhfer and Knipp (17) and Bowers et al. (10) in that confidence alone, without the need for actual answering, may fail to be a good measure of knowledge.” (2).

Neither Delores Knipp nor I have advocated that knowledge surveys or confidence measures should be used alone or should replace evaluative tools like graded tests, assignments, and projects. In our teaching, we employ all of these together with knowledge survey assessments. In our practice, we advocate that good evaluation and assessment employ multiple measures.

No one should fault authors Favazzo et al. for their perception of a “disagreement.” Their perception results from earlier misrepresentation of our work (see Ref. 4) and by Ebert-May’s and Weber’s editorial (I), which included:

“Evidence from Bowers et al. did not support the claim by Nuhfer and Knipp that students’ learning can be predicted by perceived self-efficacy levels; rather, their data indicated that the correlation between student confidence and final grades is negligible.”

We made no “claim” that anyone could “predict” direct measures of learning from either knowledge survey self-assessments or “self-efficacy (sic).” Self-efficacy refers to confidence in one’s ability to acquire the capacity for
competence through future preparation. Knowledge surveys do not record self-efficacy; they record self-assessed confidence to meet immediate challenges with present ability. In Nuhfer and Knipp (3), we reported observing that self-assessed competency appeared to increase generally with demonstrated competency, an observation that Favazzo et al. accurately interpreted and referenced in proper context. We clarify our positions here to prevent further perpetuating erroneous attributions to us that were never ours.

Favazzo et al. approach assessment in a way that collects relevant data while promoting cognitive learning and development of students’ metacognitive skills. As their database grows, additional useful insights should appear. I thank them for their contribution and recommend that others try their approach.

Sincerely,

Dr. Edward B. Nuhfer
Professor of Geology
Director of Faculty Development and Director of Educational Assessment
California State Universities (retired)

REFERENCES


FIGURE 1. Mean item-by-item comparisons of collective (N = 1,187) demonstrated competency of college students and professors through Science Literacy Concept Inventory (SLCI) scores (Reliability (R) = 0.84) and collective self-assessed competency through Knowledge Survey of the SLCI (KSSLCI) self-assessment ratings (R = 0.93).