Assessing the effectiveness of faculty development programs

Reviewed here:


The study:

This article describes a complex study designed to assess the efficacy of two faculty development programs, one extended three-year program and one intensive five-day summer workshop, which were intended to help faculty members learn about, design, and implement active, student-centered teaching techniques. The authors note that faculty development programs are typically evaluated through participant self-reports; the unique contribution of this study is to add classroom observations to the data-collection methods used in such evaluations.

The observations took the form of videotapes of class sessions, produced by faculty members themselves (N=77) in semesters after their completion of a faculty development program. “In many cases... the camera was fixed in place, focused on the instructor” (552). Faculty submitted at most two videotapes per class taught, and the videos were rated using the RTOP (Reformed Teaching Observation Protocol), a valid and reliable instrument which is designed to assess the degree to which a class is taught using active learning techniques. The study did not collect or use baseline data.

The study also used faculty self-report data derived from several surveys. On these surveys, faculty indicated that they had gained significantly in knowledge of and firsthand experience with various aspects of active, student-centered learning:

“At the end of PD, 89%... stated that they implemented reforms... that stemmed from their participation in PD activities. A majority of... faculty reported use of specific inquiry-based and learner-centered teaching practices after PD in each class period, weekly, or monthly.” (554)

These self-reports, the authors claim, are in tension with observational data:

“In contrast with the self-reported data, observations of faculty classrooms indicated that a majority of faculty (75%) implemented a lecture-based, teacher-centered pedagogy” (554-555).

Problems with the study:

- Conceptual confusion: the authors fail to distinguish between what is happening at a particular time, and change over time. At several points, they critique the idea that PD programs help faculty to change to more student-centered methods of teaching (see quote from p. 554 above; also p. 557). But the study contains no
baseline data, so the authors have no way to assess whether change has occurred. Faculty could easily have moved from a purely lecture-based approach to teaching to implementing a few student-centered activities per semester, and this study’s methodology would have no way of confirming or disconfirming such a change.

- **Apples and oranges:** the authors assert that self-report and observational data “contrast” with one another, but they are comparing very different quantities. Consider the first quote above. Faculty are said to report the use of “learner-centered teaching practices” at least monthly. And consider the source of the observational data – videotapes from at most two class periods per course in the study. This is a very small sample, given that a typical college course meets several dozen times during the term. Active learning may just have been missed. Both the observational and self-report data could be accurate; there is simply no tension between them.

Another way to describe this problem is as a lack of necessary power. The self-reported event, namely at least monthly faculty use of student-centered teaching techniques, may itself have happened in a small percentage of class sessions. It is known that the smaller the effect one is trying to detect, the more power one needs in order to do so. But the sample of classes observed in this study is quite small compared to the total number of classes, which means that the power of the study is low relative to the potential effect.

- **Coarseness of measurement:** further, the observational data reported in this study use RTOP “categories”, which are 5 broad classes of mean RTOP scores, themselves obtained by summing five subscales of questions. Categories I and II represent “teacher-centered classrooms”. Now the authors’ main point is that “a majority of faculty (75%) implemented a lecture-based, teacher-centered pedagogy”, which must mean that 75% of videotapes fell into categories I and II. But then, unless 75% of class sessions observed contained nothing but instructor lecture for the entire class period, the use of RTOP categories conceals what may well be a large amount of variation between observations, and some of these classes might contain student-centered activities consistent with faculty self-reports. Again, there seems to be no necessary contrast between the observational and self-report data.

- **Measurement quality:** the “observations” were extremely limited in that they appear typically to have consisted of a single, fixed camera pointed at the instructor. It is possible that such a camera would fail to capture student-centered learning techniques in a way that would be visible to a rater. It is also possible that the unblinking eye of a fixed camera could itself affect how a class is taught, perhaps by encouraging the instructor to remain in view of the camera rather than circulating among students. Additionally, the choice of what and when to measure was left to the instructor, potentially biasing the selection toward or against the aim. Thus, to some degree, what the authors call “observational” data is partially also “self-reported” data.

- **Research design:** while the self-reported data collected from the surveys follow an acceptable protocol for isolating the impact of the professional development programs (e.g. pretest -> intervention -> posttest), all observational data is collected after instructors have completed the program (e.g. intervention -> pretest -> posttest). Any significant differences or, as in the case of this study, lack thereof cannot be attributed to the professional development programs themselves as no baseline information regarding how the instructors taught their courses before the professional development intervention is available.

**Alternative methods:**

The professional development (PD) mentioned in the workshop did not involve personal consultation, peer mentoring, or learning communities beyond the dates of the workshop. For most participants, the workshop was held outside of the parent institution and thus not aligned specifically with the goals of the department or college with additional incentives. Thus the definition of PD in this case is very specific and somewhat unusual. For an alternative, albeit

Prepared by Paul Baepler, D. Christopher Brooks, and J.D. Walker, Research and Evaluation Team, Office of Information Technology (http://www.oit.umn.edu/research-evaluation/), 9/12/11, jdwalker@umn.edu.