MEMORANDUM  THE UNIVERSITY AT ALBANY

TO: Leonard Lerman
FROM: David Martin
DATE: June 10, 1980
SUBJECT: Your May 21 Memo on Teaching Evaluations

Thank you for your thoughts on the Educational Policy Council’s recent recommendations on the evaluation of teaching. You raise important issues.

Let me assure you that neither the Tompkins Committee nor the EPC was unmindful of the body of research to which you allude. Indeed, the Tompkins Committee members examined several literature reviews by acknowledged scholars in that area. The EPC relied primarily on one of these summaries, a copy of which is attached. It may be one of the best; and, to my knowledge, it is the most recent. The article reflects reasonably well EPC’s views and responds more thoroughly than I could to your concern about student ratings as valid indicators of teaching effectiveness and about the magnitude of the correlations produced by most studies. For that reason, I will let the article speak for me. I think you will see McKeachie’s influence in the structure, items and instructions on the form recommended by EPC and the Tompkins Committee.

I hope it was clear in my memo that the EPC has recommended an experiment with the proposed form—a pilot study. The purpose of that study (to last at least a year or more with volunteering departments) is to assess the psychometric properties of the items: How reliable are they? With what other measures do they correlate and how strongly? What is the underlying structure of the items? It is not our intention to implement use of the form on any comprehensive basis until we are fully satisfied that it performs at acceptable levels of reliability and validity. The President, the EPC, the Council on Promotion and Continuing Appointment and I believe, some uniform, campus-wide means for assessing teaching performance across disciplines and departments is needed if promotion, tenure and related considerations of teaching are to be made on an equitable and consistent basis. The precise and final shape of that device will not be determined until the pilot project is completed. The present experiment is clearly, however, an effort to move in that direction.

EPC, CPCPA and I expect that reports of student ratings will include, among other things, some estimate of how much error may be present in the ratings of an individual instructor. The details of how the form is to be administered, scored, analyzed and reported are yet to be determined, but it is my expectation that these activities will be carried out under the supervision of faculty members familiar with the instructional evaluation literature and appropriate statistical procedures—a sort of technical advisory group. It may even be desirable to have one of these individuals serve as a technical advisor to CPCPA and other bodies to help ensure that the ratings are interpreted correctly.

Your “fear that heavy emphasis on incorporating the results to encourage...”
Leonard Lerman
June 10, 1980
Page - 2 -

wrong decisions" is not, I think, well-founded. The EPC's recommendation states that "personnel decisions ... must be based in part on a student instructional rating form ..." (emphasis is mine). I believe (as I think you do) that some evidence of teaching ability is necessary to the evaluation of faculty members in tenure and promotion considerations. The EPC and I also believe (as do you) that such evidence is not sufficient for such decision-making.

There should be no presumption that because we are attempting to develop a systematic means for assessing teaching ability across departments that research records, recommendations of department heads and departmental and college/school bodies, and the other forms of evidence we have relied on in the past will be disregarded. The evidence on teaching will, we hope, receive more credence, but that should not be understood to mean that it will in any way replace or out-weigh other evidence of professional ability and performance. (You may be interested to know that the Tompkins Committee recommended that an instructor's portfolio also include peers' evaluations of course-related materials such as syllabi, examinations, assignments, texts, class or lab displays and so on. The EPC did not concur.)

Clearly, however, the key issue is the nature and quality (the reasonableness) of the evidence for judging teaching performance. To date, that evidence has consisted of ratings from highly variable departmental forms, grape-vine reports from students and inferences from a faculty member's performance in other activities. The proposed form to be tested should give us more systematic and comparable evidence, the psychometric nature and quality of which is known. To my mind, the latter has a stronger claim to "reasonableness," bearing in mind, of course, that it is to be only one type of evidence used.

This response has been somewhat lengthy, but your note raised important issues. I suspect we have not settled the matter, but I hope that I may have allayed at least some of your concerns. Thanks again for taking the time to share your views.