Published in the Wharton School *Almanac,* January 16, 1990

The Case for Minimum Teaching Standards

J. Scott Armstrong
The Wharton School, University of Pennsylvania, Philadelphia, PA

Author's Note: The following was sent to the Wharton faculty in November, 1989, challenging a set of proposals by the Wharton Teaching Committee.[1] The committee's proposal was presented as an "all or nothing" choice. Despite a substantial amount of support for the position stated below, the Wharton Committee recommendations were passed as originally proposed; this includes punitive measures for faculty who get low ratings (referred to below as the committee's Proposal #1). The proposals said that for tenure or promotion, a faculty member must get better than an "average" rating (3.0 on a five point scale). The vote was close. It seems likely that Proposal #1 would have been defeated had a secret ballot been conducted on this item alone. Action was not taken on any of the nine proposals in my paper, and neither of the proposals on process were accepted. Since that time, faculty from other schools have read the memo and suggested that it be reprinted in Almanac in order to gain further faculty comment. They are concerned that similar events in their schools may affect the quality of the educational environment.

We all share the desire to improve the learning environment at Wharton and to have students who will be satisfied with this environment. While most of the Wharton Teaching Committee's recommendations are consistent with these aims, I believe that recommendation #1, to "Establish Minimum Standards for Acceptable Teaching,"' will be detrimental to learning. I therefore recommend that we reject proposal #1. This letter describes how I reached this conclusion, suggests alternatives, and recommends a process for resolving tie issues.[2]

Factors Favoring Proposal #1

1. It communicates that something is being done.[3]

2. It may raise the short-term satisfaction level among students.

3. It may help to increase the efforts of those faculty who currently invest little effort.

Factors Against Proposal #1

1. Teacher ratings do not provide a useful measure of learning. Unfortunately, I had not bothered to read the expensive literature on this subject hen I was involved in the implementation of Wharton 's Teacher Evaluation system. The literature contradicted my basic assumptions cm this topic. Numerous empirical studies have been published on the; relationship between teacher ratings and learning. Some studies show positive relationships and some show negative ones. Typically the effects are small. (See Dubin and Taveggia, 1968, for a review of this literature.) One such study compared the performance of students in 11 sections of a calculus course at Stanford; participants in the lower rated courses performed significantly better on a common examination at the end of the course (Rodin and Rodin, 1972.) Another study, the "Dr. Fox study," showed that an actor with a completely meaningless script was able to deliver a one-hour lecture followed by a 30-minute discussion and to achieve high ratings (Naftulin, et al.

---

[1] A faculty committee appointed by the Dean.

[2] Some historical background may be of interest. I have been teaching at Wharton since 1968 and have taught at eight other schools in seven different countries. Shortly after arriving at Wharton, 1 convinced some MBA students that there should be a course evaluation and I worked with them to introduce the first evaluation system.

[3] Some people refer to this as the politician's syllogism: "Something must be done. This is, something. Therefore, this must be done."

1973). Finally, in a massive study using before and after measures of knowledge of economics, participant ratings of basic economic courses had no relationship to learning (Attiyeh and Lumsden, 1972).

2. On theoretical grounds, one might expect teacher ratings to harm learning. One principle of learning is that the learner should feel responsible for the learning (Armstrong, 1980; Armstrong, 1983; Condry, 1977; Tough, 1982). Teacher ratings communicate that the teacher is responsible for changing the student. Students are implicitly expected to play a passive role and then to evaluate the teacher's performance instead of their own performance.

3. Another theoretical argument is that evaluation tends to interfere with open communication. This would be expected to be detrimental to the learning process in that teachers may not provide as accurate feedback.

4. Teacher evaluations can be misused. Students who are doing poorly in a course can claim that the teacher is not competent in order to imply that the student was not at fault. (This applies not only to ratings, but also to other aspects of an environment that puts student satisfaction rather than learning as the primary goal.) This misuse of evaluations has been gaining in popularity. I understand (hearsay) that it is a popular strategy at state universities that stress teacher evaluations.

5. Unfavorable changes in course content may be expected. Learning involves frustration. It means that some prior behaviors are not optimal. (For example, tell a smoker to stop smoking and you will get a negative reaction.) To obtain good ratings, it is preferable to avoid disconfirming evidence and to reinforce what people already believe (Zelby, 1974). In my studies, I have found that experiential learning produces some "hostile alienated objectors" (Armstrong, 1980). Techniques for getting high ratings are often at odds with learning (e.g., it is a good idea to reinforce what students already believe rather than to imply that their current approach to problems is deficient; it is good to avoid topics where the research findings are not clear cut; it is desirable to meet expectations of students, such as that the teacher is responsible for their learning.) Content also tends to "pop management" and current events. As an analogy to this focus on short-term satisfaction, consider the doctor-patient relationship: A focus on current satisfaction is like giving the client an aspirin to make him feel good today, whereas there may be a more fundamental treatment needed to solve the causes of the headache.

6. Experimentation might be reduced. Teachers maybe less willing to try something new once they have gained tolerable ratings.

7. The contract is changed for faculty members. They were hired under one set of rules and the rules are now being changed.

8. It might be more difficult to attract new faculty.

9. Some faculty may perceive that the solution has been made on unfounded grounds. No empirical evidence was provided to demonstrate that this policy change on ratings would be beneficial in a situation such as ours. I would be surprised if such research exits.

10. Are legal issues involved? Given that the prime mission is learning, not entertainment, there may be legal issues involved in using the present type of teacher evaluation (Miller, 1978). Miller suggests that faculty members might sue on the basis that they are being judged on a criterion that has been shown to lack validity. (His article was directed at administrators to alert them of this danger.)

11. Might extrinsic rewards, such as paying for teachers for high ratings by students) reduce the motivation for a task that many faculty find to be intrinsically rewarding? (The research on attribution theory suggests this as a possibility.)

12. Might proposal #1 cause alienation of faculty from students? In environments that stress teacher evaluations, I have heard teachers make negative statements about students and refer to them as "the enemy." (This statement is based purely on unaided observation.)

13. Might some of the faculty react negatively to the assumptions implied by Proposal #1? The proposal is implies that teaching is a distasteful activity and it is difficult to get someone to do it. It also implies that faculty are

irresponsible people who respond primarily to punishments. These assumptions are common in the management of public universities; they have produced abysmal results.

14. Would proposal #1 have an adverse effect on Wharton as a research institution? Based on the recent study of publications, citations, and peer ratings, Wharton is ranked third among business schools. This represents our comparative advantage and we should not put it at risk. (It is difficult to see how the proposed changes in teacher evaluation could improve the research-environment.)

Some Alternatives

The following suggestions are listed in order of importance as I see them.

A. Restate the first proposal as "Establish standards for effective *learning.*" A basic problem is that we have no measures of learning, which is presumably a primary objective. The committee's recommendation #5 could be restated to aim at "learning evaluations," (a point that was raised in the last faculty meeting). It really is not that difficult to assess learning. I have used end-of-course evaluations that ask students to report their success at improving skills; the responses to this survey have often differed greatly from their ratings of me as a teacher. When learning is most effective, the learners believe that they have done it themselves. This attitude also bodes well for that individual's teaming after graduation. I have also used critical incidents surveys administered six months after the completion of the course to determine what techniques or concepts students have been able to use. Other measures include looking at the grades of students in follow-on courses and looking at the number of students who took follow-on courses. Finally, we should examine what happens to our students after they leave. Are they more effective on the job? Are they recognized as being effective? Are they able to use techniques and concepts from their Wharton education? Are they satisfied in retrospect that their learning was useful?

B. Use alternative measures of teacher effectiveness. Allow teachers to obtain evidence of their effectiveness in aiding the learning process. This would be particularly relevant for those faculty who rely on experiential learning.

C. Eliminate the cohort system. I was a strong opponent of the cohort system when it was proposed because it would harm the learning environment (at the same time, I thought it would be popular among students.) The evidence presented by the committee indicated that the cohort system has created a "teaching problem." It would be interesting to determine whether the faculty believes that the benefits of the cohort system are so large that they justify the bad effects. I do not.[4] As an alternative, it may be possible to modify the cohort system to allow students more freedom of choice.

D. Examine the student complaint system to determine whether there have been changes in the measurement system. A possible hypothesis would be that the Dean's office may encourage students to complain and that students go to the Dean with complaints rather than to the faculty member. If so, it may be useful to ask whether a more fruitful approach would be to require that students first cry to work out problems with the professor. (There have been a number of cases where the students have gone directly to the Dean's office when the faculty member was under the impression that things were going well.)

E. Design the system to promote learning rather than to avoid complaints. The latter approach may lead to a system designed to cater to the poorest and least interested learners.

F. Increase freedom in the environment. Rather than restrictive rules, we should seek ways to increase freedom of choice for the faculty and students.

G. Base recommendations on the empirical research on how people learn. An extensive literature exists on this topic.

---

[4] Faculty at, other schools have also reported detrimental effects on learning from the imposition of cohort systems. (This is hearsay.)

H. Define the problem from a systems viewpoint. Can we improve the environment for research and learning? 1n particular, could the learning environment in the second year of the MBA program be improved?

I. When extrinsic incentive are necessary, focus on rewards rather than punishments. How many of us believe that we respond better to punishment?

Suggestions on Process

I. Survey faculty anonymously on these issues. They are highly sensitive and some faculty have told me that they do not feel free to speak out about it. This is especially important for those who do not have tenure, but I have also heard it from those who do have tenure.

II. Seek consensus on this issue. Rather than rushing to a solution, I urge that we seek solutions that would be acceptable to most of us. If it does, become necessary to vote, I request that the voting be done anonymously.

Conclusions

My opposition against proposal #1 is balanced by my support for all other recommendations by the Wharton Teaching Committee with one exception. That exception is for #2d which relates to reduced teaching loads for junior faculty. This recommendation has no clear relationship to the teaching issue. It should be handled as a separate issue.[5]

I recommend passage of the Teaching Committee recommendations except for Proposal #1 and #2d. A vote for #1 is a vote against the research in this area. My prediction, if #1 is instituted, is that our learning environment will become more like that in most public schools; those environments are less conducive to research and learning.

Hopefully, the proposals will be handled separately. If not, I recommend a vote against the package.

References

Armstrong, J. S. (1980), "Teacher versus learner responsibility in management education," Wharton Marketing Department Working Paper.

Armstrong, J. S. (1983), "Learner responsibility in management education or ventures into forbidden research (with commentary)," *Interfaces,* 2 (April), 26-38.

Attiyeh, R. and K. G. Lumsden (1972), "Some modern myths in teaching economics: The U.K. experience," *American Economic Review*, 42, 429-433.

Condry, J. (1977), "Enemies of exploration: Self-initiated versus other-initiated learning," *Journal of Personality and Social Psychology,* 35, 459-477.

Dubin, R. and T. C. Taveggia (1968), *The Teaching-Learning Paradox.* Center for Advanced Study of Educational Administration, University of Oregon.

Miller, D. W. (1978), "Dangers of using student evaluations for administrative purposes*," Collegiate News and Views,* 2-3.

---

[5] An alternative argument is that it may be more equitable to provide re duced teaching loads for a specific purpose (e.g., to work on teaching skills or to pursue a particularly fruitful line of research) rather than to base it on length of service. Under the committee's proposal, those with less previous contributions to the school benefit more. Is this the message that we want to send?

Naftulin, Donald, J. E. Ware, Jr. and F. A. Donnelly (1973), "The Doctor Fox lecture: A paradigm of educational seduction," *Journal of Medical Education,* 48, 630-635.

Rodin, M. and B. Rodin, 1972, "Student evaluations of teachers," *Science,* 177, 1164-1166.

Tough, Allen (1982)*, Intentional Changes*. Chicago: Follett.

Zelby, L. W. (1974), "Student-faculty evaluation," *Science*, 183, 1267-1270.