

## Guest Comment: MIT Surrenders the Scientific Method

by James Steiger, Ph.D.

*Ed Note. James Steiger is professor of psychology at Vanderbilt University. His areas of expertise include structural equation modeling, statistical power analysis, and multivariate correlational statistics. With SI editor Patricia Hausman, he is author of Confession Without Guilt, a productivity analysis of MIT biologists prompted by the Institute's 1999 report on sex discrimination. Here he comments on MIT's recently released [follow-up reports](#) that again claim a hostile work environment for female faculty.*

Thanks to uncritical media coverage by newspapers like *The New York Times*, MIT initially reaped a whirlwind of favorable publicity from its 1999 report confessing to pervasive sex discrimination in its School of Science. However, those trained in research were not so readily impressed. Several commentators took MIT to task, finding its report sadly lacking in the hallmarks of a genuine scientific inquiry. Although it should have been embarrassed, MIT appears undeterred. It has now extended the exercise beyond the School of Science—committing more embarrassing failures of inference and scientific procedure in the process.

The latest effort takes the form of several reports. At first glance, the hundreds of pages of text and graphs may create the appearance of an exhaustive inquiry. Yet, with even a cursory reading, an undergraduate trained in research methods should be able to see the reports for what they unfortunately are: nothing more than a political agenda cloaked in the trappings of scientific inquiry.

Like so many "gender equity" studies, MIT's latest begin by assuming that an imbalance of the sexes in particular academic specialties indicates not only a problem requiring correction, but discrimination. This faulty assumption is but the first of many methodological errors, only a few of which can be discussed here.

### The Sloan School of Management Report

In introductory comments, the Dean states that, "By far the most surprising aspect of the Committee's work is its profoundly disturbing analysis of faculty experience. This analysis makes it inescapably clear that in our culture, men and women faculty with outwardly very similar careers are, in effect, working at two different Schools, and that the women are at a much less congenial and supportive Sloan than the men."

The basis for this "inescapable" conclusion? A survey of 12 "matched" faculty members—six male, six female—that examined 10 dimensions of "faculty experience." The committee interviewed six female faculty, and chose a "matched" male faculty member for each of them. It then asked each faculty member about their experiences at MIT. Raters then coded these "open-ended" interviews. Each experience dimension was coded as positive, neutral, or negative.

Male rated higher than female	40
Female rated higher than male	0
Both rated equally	17
Missing data	3

At first glance the data are striking. Of the 60 paired comparisons, the results show: To the uncritical—a group that, regrettably, appears to include the Dean himself—this is compelling evidence. But anyone with a modicum of training in survey design would ask at least three questions.

### ***How were the "matched" male faculty members selected?***

The body of the report assures readers that Appendix A contains "exact methods and coding categories." To the contrary, the appendix provides no information whatsoever as to how male and female faculty members were "matched." Was it simply by rank and years of experience? Or were measures of productivity (e.g. publications and citation counts) included? Absent matching for the latter, there can be no confidence that outcomes are not a logical consequence of differences in performance. Most productivity studies do show substantial differences between males and females, even after controlling for experience and rank—as did our analysis of senior MIT biologists. It is possible that more productive scientists will tend to be happier, as well as to find their surroundings more positive.

### ***Were raters "blind" to whether the subject was male or female?***

Astonishingly, Appendix A is once again silent. Yet, social scientists recognize that in studies of this type, coders must be unaware of the sex of the individuals they are rating. It seems clear, however, that the data were not collected by disinterested interviewers trained in survey techniques, but rather by a faculty committee with a clear ideological bias. Some may have felt a stake in the outcome as well.

Perhaps to suggest impartiality, Appendix A points out that each transcript was reviewed by both a male and a female rater. Again—and as any well-trained under-graduate would realize—this information is virtually useless. Nothing in the report indicates raters were unable to determine whether they were reviewing the transcript of a male or a female. Nor is there any indication that the male and female raters were selected in a way that would enhance the probability of impartiality.

### ***Does a reasonable basis exist to believe that men and women are equally forthcoming in response to questions about satisfaction and happiness?***

Even if male and female subjects were properly matched, and raters blind to the sex of each subject, we can have no confidence in the conclusions unless we have reason to believe that males and females are equally willing to disclose dissatisfaction. Absent such evidence, we are faced with the very real possibility that the results reflect nothing more than sex differences in the willingness to voice complaints.

In fact, there is strong evidence that the sexes generally differ in self-reporting behavior. For example, males report fewer feelings of depression than females. Suicide statistics, however, are inconsistent with self-reports—males take their own lives far more frequently. One must also consider that original interview responses were gathered by the committee itself. Perhaps the ideological biases of the committee affected the way both men and women responded. Recognizing the clear bias of the committee, males may have been guarded, while females felt encouraged to expound on all negative aspects of their careers at MIT.

### ***The School of Architecture and Planning Report***

A common problem with reports on the status of women is a tendency to use data in whichever manner is most consistent with an interpretation of pervasive bias. Figure 22 of the SAP report exemplifies this tendency. It presents data on tenure by sex. Only two outcomes are reported: "received tenure" or "left before tenure or denied tenure."

Any unbiased researcher would immediately see the problem here. "Leaving before tenure" might actually mean that an individual left the Institute after receiving an offer of tenure and a large salary increase elsewhere. It is hardly inconceivable that

women talented enough to be hired at MIT are, in fact, in high demand. My own department lost a very talented young woman in precisely this way. A year before her tenure decision, an American university offered a position not only to her, but also to her husband. Yet, writers of the SAP report do more than ignore this important distinction. Early in the report, they adopt a convention whereby those leaving voluntarily are lumped with those denied tenure—and classified as "*failures*."

Surely the report writers realize that a "failure" could actually be a star. We cannot even rule out the possibility that most of the male "failures" actually "failed" to get tenure, while most of the female "failures" received it, albeit outside of MIT.

MIT must have realized that this ambiguity could be resolved with appropriate data. Eventually, the report reveals that interviewer Christie Baxter, who is apparently female, was able to obtain "exit interview" data from a handful of females who left the Departments of Urban Studies or Architecture—but *not one male* from Architecture.

Baxter's interviews focused almost exclusively on feelings, and her findings are not reported in any systematic way. Strangely, despite MIT's determination to uncover gender bias so subtle as to be subconscious, Baxter and her committee did not grasp the obvious: that differences in response rates among men and women might have been related to her being female. Interviewer sex can be a significant factor in survey research, and in this case, males may have been less open to discussing their "negative experiences" than they would have been with an interviewer of their own sex.

For whatever reason, Baxter failed to establish whether those who left did so because of other offers. The simple fact is that without this distinction, the data

presented cannot have the meaning that MIT ascribes to them.

Nonetheless, one thing is clear. Despite several rounds of well-deserved criticism, MIT is still more interested in proving itself guilty by any means necessary than adhering to time-honored scientific procedures. It had the opportunity to avoid the type of methodological shortcomings of its first report, but chose not to. As a result, what ultimately emerges from its new reports is not bias against women, but against the scientific method itself.

*NAS Science Insights*  
*May 2002*