Collaborative Research:
Pitfalls and Benefits for the Statisticians

Dennis K.J. Lin

IBM Watson Research Center and University of Tennessee

1. Why collaborative research? All scientific researches share the common goal--making an impact. Consequently, researchers are always encouraged to do something *useful*. The term "useful" is not well-defined, however. Consider the famous quote from John Tukey that

$$\text{Practical power} = \text{power} \times \text{probability (actually used)}.$$

Many believe that it is nature to extend the above equation to

$$\text{Practical power} = \text{power} \times \text{probability (actually used)} \times \text{probability (efficiently used)}.$$

It is clear that to increase the practical power, the research work need to increase not only the power term, but also the probability (actually used) and probability (efficiently used) terms. Much attention have been focused on searching new technology to strengthen the power term in the past. It is equally important, however, to develop new methodology to increase last two "probability" terms. These parts seem to be the major issue in collaborative research, and in fact, is desperatedly needed in our statistics profession.

2. The importance of collaborative research. In the book "Breakthroughs in Statistics", Kotz and Johnson (1992) collect more than 50 original papers that were believed to be the most critical articles in various statistical areas. Table 1 below shows

certain subjects as examples. (For more details, see Kotz and Johnson, 1992.)

Subject	Author
Experimental Design	R. A. Fisher
Response Surface Methodology	G. E. P. Box
Time Series	Box and Jenkins
Life Table	D. R. Cox
Bootstrap	B. Efron
Ranking Test	F. Wilcoxon

It should be clear from such a list that all important pioneer statistical work were originated from the real world problem--a typical collaborative research.

Collaborative research provides the fundamental motivation for basic and applied research. Specifically, (a) collaborative research is indeed needed; (b) collaborative research is the future; (c) collaborative research is more likely to result pioneer work; (d) collaborative research is a low hanging fruits.

3. Nature Limitations. For young faculty in colleges to do collaborative research, there are two nature limitations: (1) Ph.D is only short term training, and (2) the reward system does not encourage faculty to do so. Let me elaborate more.

Apart from the basic course work, a Ph.D degree mainly emphasizes independent research ability and is normally reflected in the Ph.D Thesis. Such a thesis emphasizes new results on scientific methodology or fundamental theorem. As a result, topics with narrow focus are more likely to be chosen so that students can complete their degree requirements within 4-5 years. Collaborative research is too time-consuming to be considered in general. Under such a training, many fresh Ph.D remain to have a narrow

vision on research for the first few years period.

The current reward system (more specifically the tenure system) values the number of publications, rather than the quality of the publication (mainly because quality is hard to evaluate). The tenure process normally takes place within 6-8 years from getting a Ph.D. Young faculty is thus forced to work on narrow topics that can result in fast publications. Once again, like the case of Ph.D thesis, collaborative research is too time-consuming to be considered.

4. How to begin a collaborative research. To begin collaborative research, four basic requirements must be fulfilled. (1) interest, (2) opportunity, (3) team work, and (4) marketing. The statistician must start with an interest in certain topics before the collaborative research becomes possible. Working environment is then critical. This implies the necessary contact with other disciplines. In principle, such contacts can be easily found in Business school, if the statistician is willing to do collaborative research. As a statistician, we normally need to make our relatively newly developed techniques known to the schools, i.e., *marketing* statistics to other fields. Once initiated, the team work skill plays an important role to ensure the successfulness of the joint work.

On the other hand, we may consider encouraging collaborative research work from our curriculum (graduate or undergraduate level) by the following possibilities. (1) Internship: students are required, as part of the course work, to visit and possibly involve with real problems from BIG (Business, Industry or Government). The Statistics Master program at University of Tennessee has been very successful in such internships. (2) Consulting Course: for a large number of graduate students, internship seems

difficult. For these cases, a consulting course is a very important way for the students to have contact with real problems through the faculty's effort. The Statistics program at University of Wisconsin, for example, is well known for having such a course. (3) Lab Hours: one simplest way to make student get use to other disciplines is required them to take one semester of lab hours from other departments-- engineering, biology, etc. Students will learn how to communicate with other fields. (4) Forming Partnerships: We need to make an effort to form partnerships with other fields. The technical details may vary from one school to another, but the basic concept is the same.

Regarding the current reward system mentioned previously, I would suggest (1) only the best five (say) publications will count. These five papers will be carefully reviewed via experts. This way, faculty is more likely to concentrate to produce important work. And (2) require at least one joint paper from other disciplines, rather than focusing on the solo papers.

5. Successful Example for Collaborative Research: UTK MDC. The Management Development Center (MDC) at University of Tennessee was founded at 1972 with private money and initiative. The initial effort was a four week course in educational development programs. The growth in the late 70's was through short seminars. In 1981, its strategic shift was to longer, more developmentally oriented programs. The center has grown. In 1977 MDC had four staff, 10 programs, and $100,000 in revenues. Currently MDC has 1,500 participants, 50 programs and $6,000,000 in revenues.

MDC essentially is an integrated learning center. It consists of partnerships between business and engineering colleges. It also provides partnerships with select external partners in a co-learning process, to rethink existing managerial and

organizational paradigms with industry as a co-learner. It has no doubt impacted research, the faculty as well as curriculum. I myself have fortunately been involved with MDC, learning real world experiences, consulting techniques, and many research idea.

Discussion

- Reward System: As mentioned by the speaker, the current reward system seems to encourage faculty to have more publications. This inevitably forces young faculty focus on narrow but fast/quick publication work, rather than decent development on new methodology. The later generally takes a much longer time. Collaborative research is too time consuming in this case. Unless the current reward system changes, collaborative research will only occur at tenured level.
- Essence of Research: The essence of research is not to publish another paper, but to make differences on the way people do and the way they think. The speaker, for example, has just suggested one possible remedy for the reward system.
- Environmental: How will the size of the campus, the size of statistics faculty impact the collaborative research? Further study in this direction seems necessary.
- Statistics in B-school: Statistics in B-school is most likely under the survival mood. How to reform such a situation?
- Break down bias on each individual.