SAS® AND THE UNIVERSITIES
Rudolf J. Freund, Texas A&M University

1. BACKGROUND

For many years SAS Institute and the Universities have enjoyed a symbiotic relationship. Faculty and graduate students at Universities have found the SAS System to be an excellent research tool. In addition, the simplicity of the SAS language has enabled the use of the system as a valuable aid for the teaching of statistics. Not only does this provide students with a better understanding of a subject which is often erroneously perceived as an exercise in plugging numbers into formulas, but it also increases the pool of students who will fondly remember the SAS System as a (relatively) painless way to use computers.

SAS Institute has benefitted from this relationship in that a very large number of students have taken their knowledge and appreciation of the SAS System into the workplace where they have been instrumental in promoting the use of this system in industry and academia. Also we must not forget that SAS Institute owes its existence to a small number of Land Grant university statisticians who were instrumental in providing the funding required to establish the beginnings of the SAS (formerly Statistical Analysis System) as we know it now. Of course, the universities have benefitted by being able to use the SAS System as the primary tool for the statistical analyses of their research data.

However, it now appears that the use of the SAS System as a teaching aid is not keeping up with the ever increasing use of computers in the classroom or even with the use of computers at universities in general.

It is my belief that these trends threaten this symbiotic relationship, and although these trends are not a major threat to either SAS Institute or to the universities, they may have negative consequences. In this discussion I will briefly outline the reasons for the emergence of the relationship, factors that currently affect it, and some ideas that may help to reverse its decline.

I will concentrate on three aspects that affect this relationship:
A. the SAS System
B. documentation
C. communication.

Since most of the SAS usage at universities revolves around statistics, the main focus of these comments will revolve around this type of application. Furthermore, the discussion will emphasize the role of the SAS System in students' educational experiences at both undergraduate and graduate levels.

2. THE PAST

This symbiotic relationship was a natural outgrowth of a number of favorable factors:

A. A truly remarkable software product for data handling, statistical analysis, and report writing. The SAS System was and still is a flexible and easy to use tool for data management coupled with a comprehensive line of statistical analysis procedures.

B. A reasonably complete, compact and low priced Users' Guide. As is the case for much software documentation, it was primarily a reference work and required some additional introductory instruction. However, considering the price, it was not an unreasonable reference work for student use.

C. Excellent lines of communication between the users and the software developers at the SAS Institute.

3. CURRENT STATUS

As the world changes, so has this relationship between SAS Institute and the universities.

SAS Institute has become a large corporation with many divisions which are responsible for a number of different 'products'. An ever increasing proportion of these products cater to data management and report writing and other industrial applications, with statistics and the use of the SAS System at universities becoming an ever decreasing proportion of SAS usage.

At universities (as elsewhere) the use of computers has mushroomed, as has the availability of software usable in both teaching and research efforts. In statistical education, some of this software is supplied as a supplement for a particular textbook, and usually consists of a very limited and inflexible package having very little usefulness beyond its use with that textbook. Alternatively, use is made of one of the many more general software packages.

A. The SAS System remains, as it has been: a comprehensive and flexible data analysis, data handling, and report writing system. Since statistical data analyses are the backbone of much of the research performed at universities, the use of the SAS System for such purposes continues at a high rate.

The role of the SAS System in the education of students is not as clear. As computers are increasingly used as a teaching tool, some of the requirements for the suitability of software for use in statistics courses become
increasingly clear:

(i) Ease of use (User Friendly?!).
(ii) Flexibility, that is, allowing the user
to perform a wide variety of statistical
methods.
(iii) Good graphics.
(iv) Modest resources requirements so that
it can be effectively used by a large
number of students.
(v) Interactivity.
(vi) Ability to easily use data sets from
externally supplied files.
(vii) Future usefulness for students.
(viii) Cost.

The SAS System obviously shines for
requirements ii, vi and vii. In fact, one of
my strongest arguments for using the SAS System
is that knowledge of this system will benefit
students in many other courses as well as in
their jobs. Software especially written for
instructional purposes is often easy to use,
has quick execution and usually requires
minimum computer resources. However, the use
of such software leaves students with very
little they can use, either in other academic
endeavors or in their careers.

The ability of the SAS System to meet the
other requirements is not as clear. Flexibility
and the need to access a variety of data files
(requirement iii) does require more programming
which diminishes the user friendliness and
requires more resources (i and iv). For
example, in many systems you simply request a
regression and get a fixed set of results
including one or more standard plots. However,
in some ways this is much like the mindless
plugging of numbers into formulas. On the
other hand, in the SAS System you need one set
of statements for doing the regression, then
specify what variables are needed for plotting
and finally implement a plot. And if a printer plot is not adequate, implementing SASGRAPH is
not all that simple.

Interactivity is indeed wonderful. However,
the more flexible a system is, the more
difficult becomes the implementation of inter­
activity, not to mention user friendliness. I
recently tried using a competing system, and
even though it was not very flexible, it was
longing for a command structure after wading
through the various menus and submenus.

Interactivity also requires more resources,
especially with a more flexible system. On the
other hand, the batch oriented mainframe
(especially the 'autobatch') version 5, can be
used by hundreds of students without tying up
the system. Of course, the ever increased
power and reduced cost of computer hardware may
tend to alleviate the resources problem.

The cost of the SAS System is certainly
not a barrier to its use for instructional
purposes, especially for mainframe users, since
in this case the system is also used for
research and other purposes. However, the
piecemeal pricing of the PC version (even with
relatively low prices) does discriminate against
those needing to use the SAS System for
statistics.

All in all, the SAS System certainly has
many features that make it useful for instruc­
tional purposes, but not without some reserva­
tions.

B. Documentation. The current status of SAS
System documentation is a major stumbling block
to the effective usage of the SAS System for
teaching. The primary Version 5 documentation
consists of the introductory Guide and the two
volume Users' Guides. This is equivalent to
providing a first grade reader and a Webster's
 unabridged dictionary for someone wanting to
learn how to use the English language! Obviously
this is a very serious matter for those wanting
to use the SAS System for teaching statistics:
the introductory Guide is too limited and
students certainly can neither use nor
reasonably be asked to buy the dictionary. And,
unfortunately, neither SAS publications nor
anyone else has provided any reasonable
alternatives. The SAS Series in statistical
applications are useful primarily for more
advanced courses and, of course, research.

The situation for Version 6 is, if
anything, worse. At this time four or five
guides are needed requiring a major financial
investment. And invariably what you want is in
another volume! Remember also that the computer
instruction in a statistics course is supple­
mental: the students have already invested in
the primary text.

There is in preparation a new guide on the
order of the SPSS® (1) Statistics guide. This
guide is a combination text in statistics and
SAS and may be useful for the casual user
(technician) but may not be suitable for
instruction where the instructor will want to
rely on the primary statistics text to teach
statistics.

C. Communications. The obvious need to place
some restrictions on the access of individual
users to the SAS Institute staff has prompted the
use of a single SAS installation represen­
tative. This representative is normally
housed in the computing center and usually has
neither interest in nor knowledge of statistics.
Such a representative is obviously useless for
consulting and troubleshooting in statistical
applications. Also, in some cases the represen­
tative is not very available to many users.

4. WHAT IS NEEDED?

A. The SAS System serves most purposes
reasonably well. A few suggestions may merit
consideration:

(i) A 'mini-SAS' for microcomputers.
know this has been steadfastly opposed by SAS Institute. Obviously, such a system would benefit those who want to use the SAS System for teaching, but I also believe that such a system could increase the number of 'casual' users of the SAS System who could eventually become 'serious' users and thus broaden the market for the system. Admittedly, the existence of a 'mini-SAS' could create problems both for SAS Institute and users and the increasing capacities of PC's tend to diminish the need for such a system. Nevertheless, it is an idea that should be considered.

(ii) The separation of Base-SAS, SAS-STAT, and IML for microcomputers into separate 'products' is an insult to statistics and statisticians. SAS Institute should remember its heritage!

(iii) I hope SASGRAPH for microcomputers is easier to use. Good graphics are vital for teaching statistics.

(iv) SAS Institute should consider publishing a set of Macro procedures that implement relatively complete analyses for a number of reasonably simple statistical analyses. These could be used to teach the basics of the system which could then be expanded by teaching how to make the analyses more comprehensive by using the power and flexibility of the SAS System.

(v) Continue to maintain an 'Autobatch' version of SAS for mainframes.

B. Documentation. The most pressing need is for a one volume, general purpose Users' Guide. Once upon a time SAS had a single volume Users' Guide that worked quite well. Admittedly the SAS System has grown, but with the current (and presumably growing) set of specialized Guides, it should be possible to provide most users about 95% of what they need in a single compact volume of 400 pages or less. And lest the ever dollar hungry publications division be afraid that the sales of the dictionaries be diminished, I believe this decrease will be more than compensated by large sales of the single volume Guide.

Another need in the publications area are SAS based materials to be used in education. It is fair to state that millions of potential SAS users are currently enrolled in our nations universities. Exposure to the SAS System would be very beneficial to many of these students and would certainly not hurt SAS! However, such exposure requires instructional materials that are not available, either through SAS publications or anywhere else. Since many students take some sort of elementary statistics course, a supplement that can be used in such a course would be quite useful. I have noted that many of these courses are using Minitab® and/or software especially produced for use with a specific text. Many textbooks do use SAS (or SPSS, Minitab, or combination) outputs to illustrate statistical methods, but this is not the same as using these products in the instructional process. When using such texts, some sort of supplemental material will be required, and since SAS prides itself on an extensive publications list, it should seriously consider such a publication.

Similar materials should also be considered for courses which use some of the other applications of the SAS System, particularly in business, operations research, and quality control courses.

C. Communications. SAS Institute should seriously consider having specialized installation representatives. It seems that as long as SAS offers the system as a set of separate 'products' there could be a representative for each product. It would, for example, be very useful to have a SAS statistics representative who could not only diagnose a problem with a statistics procedure, but also then know who at SAS Institute could best solve the problem. This need is, of course, not unique to universities.

5. SUMMARY.

SAS Institute has provided us a wonderful tool for use in our profession. Because we have found it so useful, we would like to share it with our students. We believe that this sharing is beneficial to all: faculty, students, and SAS Institute. Unfortunately, as this sharing becomes increasingly useful for students and faculty, the policies of the SAS Institute make this sharing increasingly difficult. As we have noted before, SAS can certainly flourish without us, and we can presumably find other software to fill our needs. But, is this what we really want?