VIGRE in Statistics at Carnegie Mellon

Rob Kass

Department of Statistics
Carnegie Mellon University

May 9, 2003
Goals

- To provide students and postdoctoral fellows with
  - a solid foundation of mathematical statistics and statistical methods, and
  - serious cross-disciplinary experience
while emphasizing
  - adaptability and breadth, together with
  - good communication skills, including an appreciation for the
    cognitive needs of the learner.

- To help increase the number of U.S. citizens, and especially
  women and minority students, who receive advanced training in
  Statistics.
Organizing Principles

At all levels:

- Core statistical theory and methods are fundamental.
- The practice of statistics requires special skills.
- It is worth taking account of knowledge about learning and expertise.
- Much the same communication skills are needed for cross-disciplinary work and teaching.
OUTLINE

1. Context
   - Statistics and Carnegie Mellon
   - Learning and Expertise
2. Training via Statistical Practice
3. Environment and Diversity
Statistics and Carnegie Mellon

Statistics as a field is booming, and the essential core discipline (theory and methodology) must be redefined and reinforced.

Historical strengths of Carnegie Mellon have suited Statistics extremely well:
- Human decision-making, which encouraged growth in Bayesian statistics;
- Computing, which has increased in importance (relative to mathematics);
- Cross-disciplinary research (continued on next slide).

All of these are now widely recognized as crucial, though none was at the time we embarked down these paths.
We Collaborate Extensively

- Astrophysics, neuroscience and brain imaging, genetics and molecular biology, environmental and computer engineering, data mining, education and psychometrics, finance, real-time computing, psychiatry, social science, government, and public policy.

- All of these areas are open to Vertically-integrated teams of trainees. We have a variety of such teams.
Our Research Environment

- 18 Full-time Regular Faculty:
  - 10 Professors, 2 Associate Professors, 2 Assistant Professors,
  - 3 Research Scientists, 1 Lecturer.
- Professors: middle-aged and eminent.
- Over 40% of our publications are cross-disciplinary.
Evidence about Learning and Expertise

Comes from two sources:

- Laboratory experiments on humans. (What helps or hurts in learning simplified material?)
- Cognitive theories, implemented as computer programs, that can mimic aspects of human behavior.

As opposed to much speculation in the educational literature, there are scientifically solid results in lab settings. It not an unreasonable leap of faith to expect these results to apply to knowledge and skill acquisition generally.
Learning and Expertise—Two Examples

EXAMPLE 1: Experiment by Bower et al. (1969) on word recall found that accompanying the words with an easily-understood organizing structure greatly improved recall.

From J.A. Anderson, *Cognitive Psychology and Its Implications*:

- When information is committed to memory it is often elaborated with additional information [which can] facilitate recall by providing additional retrieval paths and by permitting recall by inference and reconstruction.

- Schemas are a major mechanism for elaborating material during study, and they are a major mechanism for reconstructing memories at test. Recall will be distorted to fit the schemas that a subject has.
Learning and Expertise (Continued)

EXAMPLE 2: Chi et al. (1981) found that physicists classified elementary physics problems according to principles (such as Newton’s second law) while novices did not.

Anderson:

Improvement [in skill learning can come] about [when] people learn the optimal way to organize their problem solving for a particular domain.

These general results may be applied to the way we help people learn statistical concepts, and develop statistical expertise.
What Are We Expecting of Students?

“Getting it” in Statistics requires integration of a nontrivial conceptual overview.

A key component is the link between theory and practice.
Conclusions

Experiments or Observations

Unobserved Mechanisms

Key Features

DATA

regularity  variability

Algorithms

EDA

Conclusions

REAL WORLD

THEORETICAL WORLD

Rules of Probability

Probability Models

parameters  noise

random variables

Formal Statistical Methods

Conclusions

Conclusions
How Should We Help Students?

Goal: Students should become proficient in understanding how statistical methods can be informative in “real-world” data analysis settings.
Training via Statistical Practice:
We Call This The $ABA^{-1}$ Process

Training links research and education.

$A$: scientific question $\rightarrow$ statistical problem;

$B$: solve problem;

$A^{-1}$: translate results back to relevant community.

- Theory and methods [$B$] must be connected to real problems [$A$].
- Students in elementary statistics need to internalize this connection.
- Advanced graduate students need to deepen their understanding of the process.
- Teaching to collaborators [facilitating $A$ and $A^{-1}$] or beginning students should follow basic principles.
Seven Principles of Teaching
(S. Ambrose, Carnegie Mellon)

1. Prior knowledge is the basis for building new knowledge.
2. How knowledge is encoded and organized determines its access and use.
3. Active engagement can promote deeper learning.
4. Feedback that is timely, frequent, and constructive is crucial to learning and the development of expertise.
5. To develop proficiency, learners must acquire the skills of selecting, monitoring, evaluating, and adjusting their learning strategies.
for Undergraduates—Highlights

- **Statistical Reasoning**, with StatTutor
- Statistical Theory
  Streamlined, then linked to practice.
- **Data Analysis**
  Year-long undergraduate course, the second semester of which is a project course.
for Undergraduates—Highlights (Continued)

Summer Undergraduate Research Experience in Statistics

- Small research teams, with supervision;
- “Perspectives on Statistics” seminars;
- Includes interns from Project IMHOTEP at Morehouse College;
- Will be involved with Center for Minority Health at University of Pittsburgh.
$ABA^{-1}$ for Graduate Students

**Perspectives on Statistical Practice and Instruction**
Fall-semester course for first-year graduate students.
- The $ABA^{-1}$ process.
- Panel on cross-disciplinary research.
- The writing process; feedback on writing.
- The learning process; dealing with first-year undergrads; holding office hours.
- Teaching statistical reasoning; feedback on mini-lectures; feedback on interactions with students.

**Statistical Practice**
Spring-semester project course for first-year graduate students, with external collaborator (“client”) and oral and written presentations.
ABA$^{-1}$ for Graduate Students (Continued)

- Year-long project course culminating in preliminary exam for Ph.D. Faculty advisors: one external, one internal.
- Extensive training in teaching.
  - Introductory lectures on teaching and learning during our orientation program.
  - Perspectives course (described above).
  - Cross-disiplinary training, with mentoring (described above).
  - TA activities.
  - Associate Instructor activities: significant responsibility for a large undergraduate course.
  - Departmental teaching teas; seminars from our University Teaching Center.
  - Teaching during summer, with discussion and some feedback.
Graduate Student Training—Additional Remarks

- We try to maintain a supportive environment.
- We include a mentoring seminar.
- We have streamlined our graduate program.
- Minority recruitment remains a serious issue. We expect to use summer program as “foot in the door,” for networking. We are developing a plan to solidify preparation of entering graduate students.
\( ABA^{-1} \) for Postdocs

- Visiting Assistant Professors; teach 1 course per semester (sometimes less); co-teaching when possible.
- Mentored research projects.
- Training in teaching.
  - Orientation.
  - Departmental teaching teas; visitation and feedback by University Teaching Center.
  - Mentoring discussions.
- Mentoring seminar.
Summary

- Adaptability and breadth:
  - We describe the goal in terms of $AB A^{-1}$;
  - We concentrate much effort on the $A$ and $A^{-1}$ components, including consideration of experimental results from cognitive science;
  - Our discipline is defined by $B$; we keep it solid, but streamline where possible.

- Effective communication:
  - We require many interactions and presentations as part of $A^{-1}$;
  - We supervise development of teaching skills.
Summary (Continued)

- In addition to emphasizing cross-disciplinary research, we attempt to maintain a cohesive and encouraging workplace.

- Increasing the number of U.S. graduate students in Statistics is difficult. Our efforts are aimed at both attracting and retaining students.