DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF TORONTO

CSC 428F/2514F

HUMAN-COMPUTER INTERACTION

Lecture 18

EXAMPLES OF HCI RESEARCH AND EVALUATION IN INTERACTIVE SYSTEM DEVELOPMENT

Ronald Baecker
Professor of Computer Science,
Electrical and Computer Engineering, and Management
University of Toronto

Copyright © 1991-7, Ronald Baecker.
All rights reserved.

18.1 HCI research and evaluation in the real world

Human-Computer Interface Design: Success Cases, Emerging Methods, and Real-World Context. (Rudisill, et al., 1996)

Examples discussed

- The Xerox Star (BGBG, Case Study B)
- DEC Rally Version 2 (4th generation application generator)
- The ITS user interface development system (Lect. 14)
- Freestyle multimedia office communications system (this lecture 18)
- Project Ernestine (GOMS modelling of a toll-&-assistance operator workstation (Lectures 19-21)

We'll look at 2 success cases where research and evaluation played an important role

- The SuperBook electronic book system (18.2-18.5)
- Freestyle office communications system(18.6-18.8)

18.2 Hypertext, text, and reading on screens & paper

Questions for class:

What is hypertext?

What are the advantages and disadvantages of hypertext compare to normal linear text?

How does reading on screens compare to reading on paper?

Why might reading on screens be more difficult?

Why might reading on screens be more difficult?
- Perceptual difficulties — poor image quality, justification
- Cognitive issues — lack of organization & navigation aids
- Affective issues — comfort and preferences

18.3 The theoretical bases of SuperBook

Research studies on the psychology of information access by humans, mostly carried out at Bellcore (Figure 18.1)

Results helped suggest new ways to support indexing, querying, navigation, and display of electronic text

Figure 18.1. History of Bellcore research and development on text retrieval and browsing (Landauer, BGBG, p. 664)

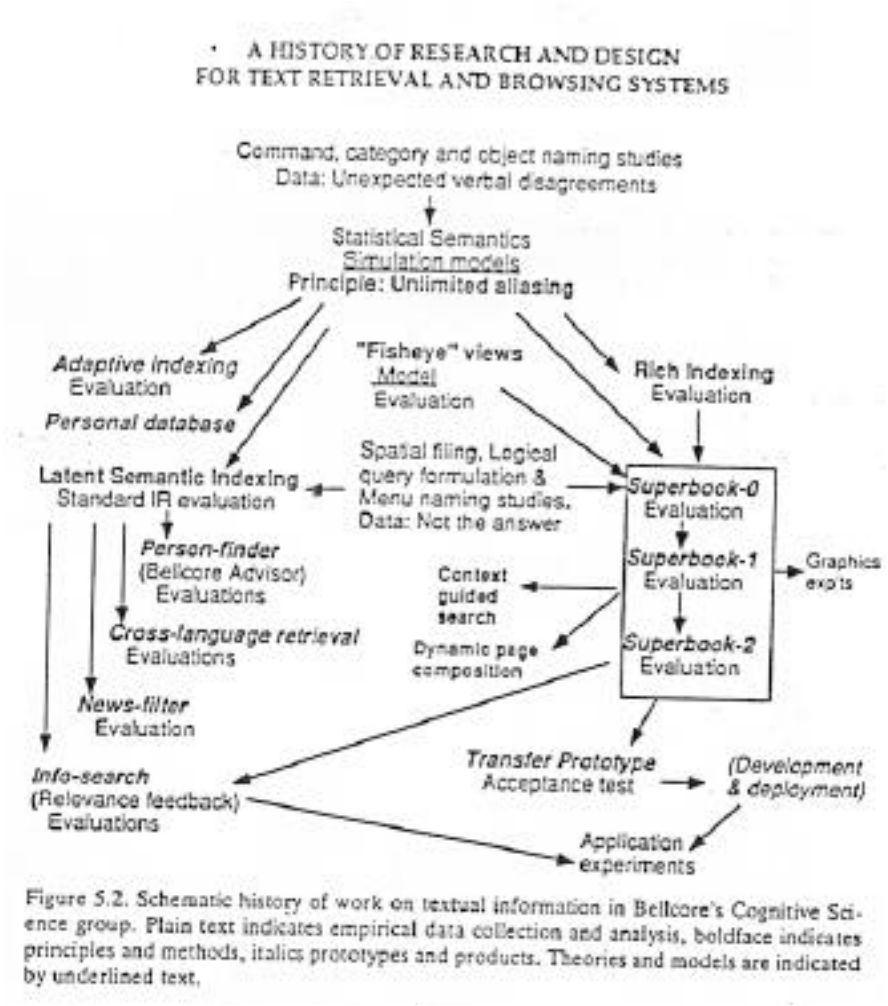

Figure 5.2. Schematic history of work on textual information in Bellcore's Cognitive Science group. Plain text indicates empirical data collection and analysis, boldface indicates principles and methods, italics prototypes and products. Theories and models are indicated by underlined text.

The vocabulary mismatch problem

People call the same thing by many different names

100 people often suggest as many as 30 different names

Most frequently used term usually used by less than 1/3 of the people

If someone else names something, and you try to guess the name, and keep trying, still less then 50:50 chance you will come up with the “proper” term

One solution: *unlimited aliasing (rich indexing)*

Give each object large number of names

Bad effect: you retrieve 2 or 3 times as much “junk”

Good effect: you retrieve 5 X as much of stuff you want

Another solution: adaptive indexing
- Assume you try 3 words that don't yield the desired result before you succeed with a 4th
- System then asks you if each of these 3 terms should be added to the list of index terms for that item

The navigation problem
- Feeling of disorientation, sense of being lost
- No clear idea of
  - Where am I?
  - Where have I been?
  - Where am I going?

One solution: Fisheye views — Display lots of detail where you are looking, successively less detail the further out you go, assuming where you are looking maximizes what Furnas calls a “degree of interest” function

Ideas that seem not to work so well — Boolean searches, complex hierarchical menus, spatial layout of menus

18.4 The SuperBook hypertext system

SuperBook also enabled by advances in computer processing, storage, and display

Key design features of SuperBook (Figure 18.2)
- Good search tools
- Structured search feedback
- Dynamic table of contents
- Text navigation facilities
- Tailorable screen layouts
- Optimized interaction protocols

Figure 18.2. A SuperBook display (Egan et al., BGBG, p. 844)

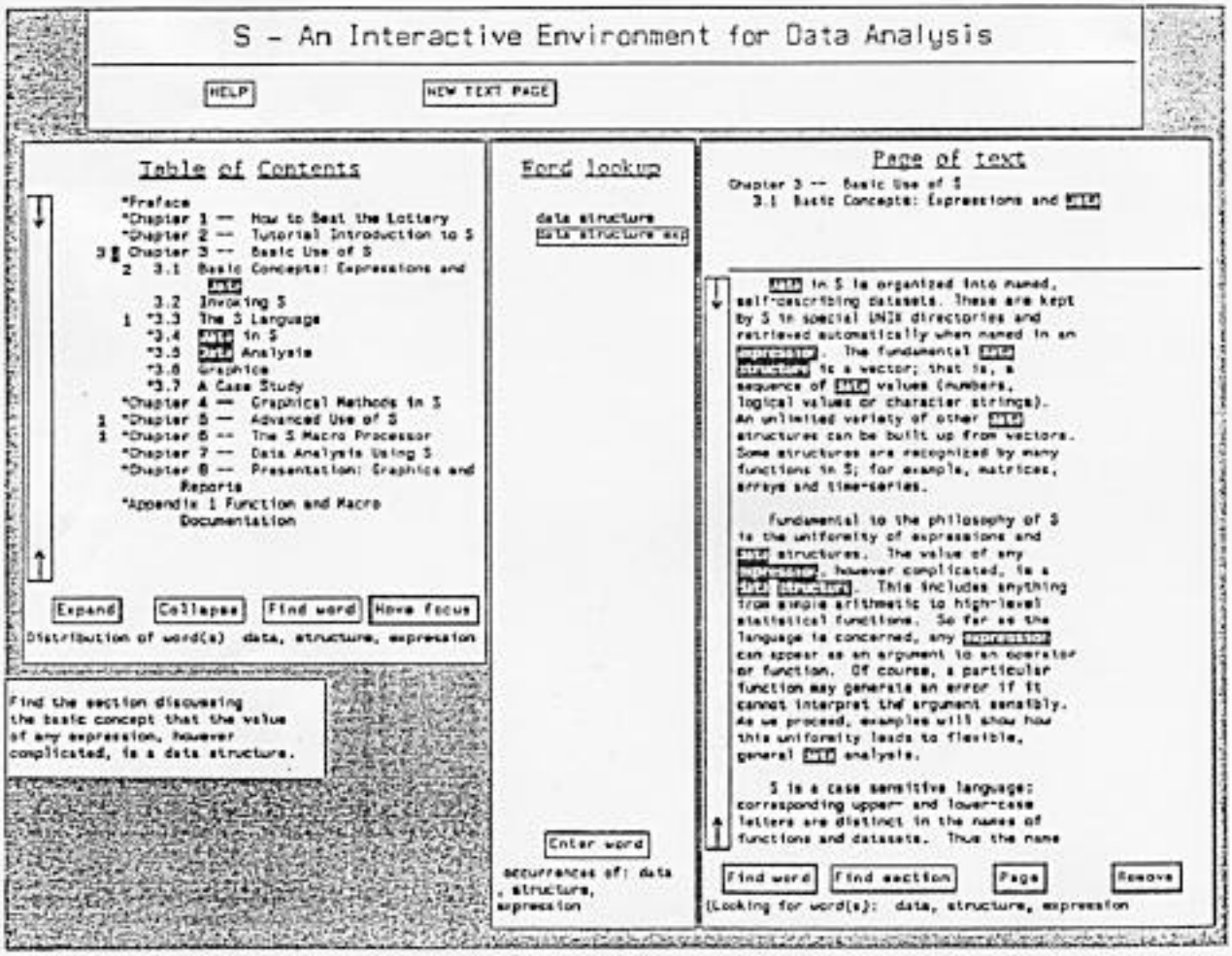

Figure 1. SuperBook display showing results of expanding the Table of Contents, looking up several terms, and displaying text containing search terms.

18.5 Success of SuperBook

Success for principled, theory-informed *formative evaluation*

User studies and experiments at Bellcore and other sites
- Measuring search time and accuracy
- Measuring quality on open-book essay questions
- Logging entire sessions for playback, also using interviews

Some of these tests (Egan et al., BGBG, 843-848) compared search time and accuracy in 4 conditions, whether or not the readers' language corresponds to that of the writer, and whether or not the reader's words are included in the document structure created by the writer

Many small improvements in system at each successive stage of formative evaluation --> dramatic results (Fig. 18.3)

Figure 18.3. Successive improvements in SuperBook (Landauer, BGBG, p. 660)

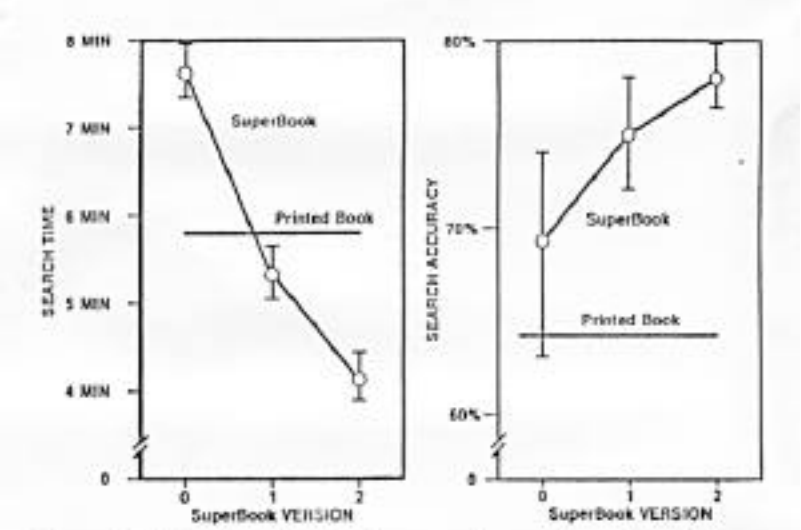

Figure 5.1. Information retrieval performance with successive versions of SuperBook. Version 0 was based on several previously formulated principles and incorporated successfully demonstrated features. Data are from an experiment in which students found answers to a variety of questions in a statistics book in its original print form or online with SuperBook. Versions 1 and 2 reflect the results of formative evaluation and redesign. About half of the time difference between versions 0 and 1 is attributable to faster code, the result to improved user–system interaction protocol.

18.6 The Freestyle office communications system

Pioneering, although commercially unsuccessful, system by Wang Laboratories

Video tape demonstration (from Siggraph Video Review 45)

Capabilities
- High-resolution graphics
- Image capture and storage
- Voice recording
- Multimedia electronic mail
- Pen markup and annotation
- Gestural interface
- Playback of synchronized voice and handwriting
- Use on networked personal computers

Design ideas
- Support use of familiar skills, e.g., writing, sketching, erasing, talking on the phone
- Benchmark the system against the physical world, requiring graphical fidelity and high performance
- Emphasize innovation
- Provide flexibility and encourage customization

Applications
- Document creation, markup, transmission, sharing, and collaboration
- Use for distributed workgroups for complex communication required for problem-solving, negotiating, planning, and design

18.7 Research and evaluation with Freestyle

Many field, experimental, and respondent strategies used during development of Freestyle (Figure 18.4)

Figure 18.4.a HCI Activities, Methods, and Results during Development of Freestyle(Francik, in Rudisill, et al., pp. 64-65)

TABLE 1 HCI Activities, Methods, and Results During Development of Freestyle

Area	Methods	Results
HARDWARE	Try to break it.	Identified performance requirements for tablet to support fluid handwriting. Selected vendor to work with on tablet design.
	Laboratory evaluation of off-the-shelf equipment and of prototype tablets and pens. Rating, ranking, observation, questionnaire.	Specified detailed design for pen and tablet. Led to working prototype hardware that was distributed to internal test site participants. Gathered initial information on hand–eye coordination difficulties in writing on a tablet while looking at the display.
	Internal sites: observation of training, follow-up interviews, ongoing observation, user roundtables.	Uncovered issues with hardware configuration on desk. Refined design and developed accessories. Made improvements for final design to improve durability. Assessed long-term ease of use, particularly hand–eye coordination, as shown in documents produced by users. Some initial behavior observed in laboratory study did not persist in actual use. Gathered contextual information on use of handwriting vs. typing.
	Customer sites: follow-up interviews and observation one month after installation.	Confirmed information on equipment configuration, use of accessories, general ease of use.
SOFTWARE	Iterative usability tests of interface and tutorial.	Collected converging information on hand–eye coordination based on initial use and detailed observation. Assessed which features required no training, minimal training, some training, or redesign. (Juxtaposed these findings with long-term findings from field sites.) Tested both new and redesigned features for usability. Refined documentation, online help, and quick reference card. Coordinated changes to product with those to documentation.
	Internal sites: preinstallation interviews, observation of training, follow-up interviews, ongoing observation, user roundtables.	Assessed fit between initial (minimal) system and existing paper process. Prioritized design change requests. Collected evidence for long-term usability of product features. Iterated design of features such as voice annotation controls and filing aids. Customization provided clues to new features. Discovered need to prompt workgroup to take full advantage of product for complex communication. This led into strategies for encouraging adoption of product.

Figure 18.4.b HCI Activities, Methods, and Results during Development of Freestyle(Francik, in Rudisill, et al., pp. 64-65)

Area	Methods	Results
	Customer sites: follow-up interviews and observation one month after installation.	Gathered broader, but less detailed, information than that received from internal sites. Conflicting requests for product changes came from many different organizations, industries, and work processes. Opportunity to see how new users do with the product; by this time, the internal sites were very experienced with it.
WORKGROUP SUPPORT	Needs-finding interviews with workgroups representing potential markets.	Uncovered new, richer scenarios of use; incorporated into marketing literature. Also used by customer site teams. Beginning of changes in how the development team thought about the product.
	Internal sites: preinstallation interviews, observation of training, follow-up interviews, ongoing observation, user roundtables.	Developed deep understanding of barriers to product introduction, use, and success. Clear evidence that the product depends on workgroup selection and configuration. Examples of pitfalls and of successful use would help customers think about similar issues in their organizations. Beta (customer) test manager became committed to workgroup analysis.
	Customer sites: work with sales support staff, preinstallation interviews, follow-up interviews, and observation one month after installation.	Encountered the same barriers to product introduction, use, and success. Provided targeted consulting to customers before and after installation. Streamlined interview techniques for rapid workgroup assessment. Attempts to teach workgroup assessment were unsuccessful; instead, summarized consulting information in a video.

Specific example of things learned

- Hardware — Size, shape, feel of pen and tablet
  (Recall “eraser crumbs” anecdote)
- Software — Integrated use of interface and tutorial
  (Testing discussed in Perkins, et al., in BGBG)
- Workgroup support — Appropriate choice of workgroup, irrespective of org. chart (Fig.'s 4, 5, BGBG, p. 893)

18.8 Freestyle as groupware

Groupware, software designed for groups or organizations, fundamentally different from single-user software

Adoption issues
- Identifying uses
  - For groups rather than individuals
  - For unstructured as well as structured processes
- Selecting workgroups
  - Typing and non-typing users
  - Numbers of systems, geographic spread, organization chart or workgroup boundaries

Product introduction strategies
- Use of consultants to advise on issues of selection of uses and workgroups
- More generally, issue of education in a new paradigm
- Critical mass problem

More on these issues in Lectures 22 and 23

18.9 References

Egan, D., Remde, J., Gomez, L., Landauer, T., Eberhardt, J., and Lochbaum, C. (1989). Formative Design-Evaluation of SuperBook. *ACM Transactions on Information Systems* 7(1), 30-57.

Egan, D., Remde, J., Landauer, T., Lochbaum, C., and Gomez, L. (1989a). Behavioral Evaluation and Analysis of a Hypertext Browser. *Proc. CHI'89,* ACM, 205-210, reprinted in BGBG, pp. 843-848

Francik, E. (1995). Rapid, Integrated Design of a Multimedia Communication System. In Rudisill, Lewis, Polson, and McKay (1995).

Francik, E., Rudman, S., Cooper, D., and Levine, S. (1991). Putting Innovation to Work: Adoption Strategies for Multimedia Communication Systems. *Communications of the ACM* 34(12), 53-63, reprinted in BGBG, pp. 886-896.

Landauer, T., Egan, D., Remde, J., Lesk, M., Lochbaum, C., and Ketchum, D. (1993). Enhancing the Usability of Text through Computer Delivery and Formative Evaluation: the SuperBook Project. In McKnight, C., Dillon, A., and Richardson, J., *Hypertext: A Psychological Perspective*. Ellis Horwood, 71-136.

Levine, S. and Ehrlich, S. (1991). The Freestyle System: A Design Perspective. In Klinger, A. (Ed.), *Human-Machine Interactive Systems,* Plenum, 3-21, reprinted in BGBG, pp. 871-880.

Perkins, R., Blatt, L., Workman, D., Ehrlich, S. (1989). Iterative Tutorial Design in the Product Development Cycle. *Proc. 33rd Annual Human Factors Society Meeting,* 268-272, reprinted in BGBG, pp. 881-885.

Landauer, T., Egan, D., Remde, J., Lesk, M., Lochbaum, C., and Ketchum, D. (1993). Enhancing the Usability of Text through Computer Delivery and Formative Evaluation: the SuperBook Project. In McKnight, C., Dillon, A., and Richardson, J., *Hypertext: A Psychological Perspective*. Ellis Horwood, 71-136.

Rudisill, M., Lewis, C., Polson, P., and McKay, T. (Eds.) (1996), *Human-Computer Interface Design: Success Cases, Emerging Methods, and Real-World Context*. Morgan Kaufmann.

Wang Laboratories (1989). Freestyle. *ACM SIGGRAPH Video Review 45*.