CHAPTER 5:
Evaluation and the User Experience

Designing the User Interface:
Strategies for Effective Human-Computer Interaction

Sixth Edition

Ben Shneiderman, Catherine Plaisant,
Maxine S. Cohen, Steven M. Jacobs, and Niklas Elmqvist
in collaboration with
Nicholas Diakopoulos

Addison Wesley
is an imprint of
PEARSON

Survey Instruments

- Written user surveys are a familiar, inexpensive and generally acceptable companion for usability tests and expert reviews

- Keys to successful surveys
  - Clear goals in advance
  - Development of focused items that help attain the goals

- Users could be asked for their subjective impressions about specific aspects of the interface such as the representation of: task domain objects and actions
  - Syntax of inputs and design of displays

Examples of the specific satisfaction scale questions:					
5.4	Messages which appear on display:	confusing	clear	1 2 3 4 5 6 7 8 9	NA
5.4.1	Instructions for commands or choice:	confusing	clear	1 2 3 4 5 6 7 8 9	NA

Survey Instruments (continued)

- Other goals would be to ascertain:
  - User background (age, gender, origins, education, income)
  - Experience with computers (specific applications or software packages, length of time, depth of knowledge)
  - Job responsibilities (decision-making influence, managerial roles, motivation)
  - Personality style (introvert or extrovert, risk taking or risk aversive, early or late adopter, systematic or opportunistic)
  - Reasons for not using an interface (inadequate services, too complex, too slow)
  - Familiarity with features (printing, macros, shortcuts, tutorials)
  - Feeling state after using an interface (confused or clear, frustrated or in-control, bored or excited)

Survey Instruments (concluded)

- Online surveys avoid the cost of printing and the extra effort needed for distribution and collection of paper forms
- Many people prefer to answer a brief survey displayed on a screen, instead of filling in and returning a printed form
  - although there is a potential bias in the sample
- A survey example is the Questionnaire for User Interaction Satisfaction (QUIS)
  - http://lap.umd.edu/quis/
- There are others, e.g. Mobile Phone Usability Questionnaire (MPUQ)

Acceptance Test

- For large implementation projects, the customer or manager usually sets objective and measurable goals for hardware and software performance
- If the completed product fails to meet these acceptance criteria, the system must be reworked until success is demonstrated
- Rather than the vague and misleading criterion of "user friendly," measurable criteria for the user interface can be established for the following:
  - Time to learn specific functions
  - Speed of task performance
  - Rate of errors by users
  - Human retention of commands over time
  - Subjective user satisfaction

Acceptance Test (concluded)

- In a large system, there may be 8 or 10 such tests to carry out on different components of the interface and with different user communities
- Once acceptance testing has been successful, there may be a period of field testing before national or international distribution

Evaluation During Active Use and Beyond

- Successful active use requires constant attention from dedicated managers, user-services personnel, and maintenance staff
- Perfection is not attainable, but percentage improvements are possible
- Interviews and focus group discussions
  - Interviews with individual users can be productive because the interviewer can pursue specific issues of concern
  - Group discussions are valuable to ascertain the universality of comments

Evaluation During Active Use and Beyond (continued)

- Continuous user-performance data logging
  - The software architecture should make it easy for system managers to collect data about:
    - The patterns of system usage
    - Speed of user performance
    - Rate of errors
    - Frequency of request for online assistance
  - A major benefit is guidance to system maintainers in optimizing performance and reducing costs for all participants

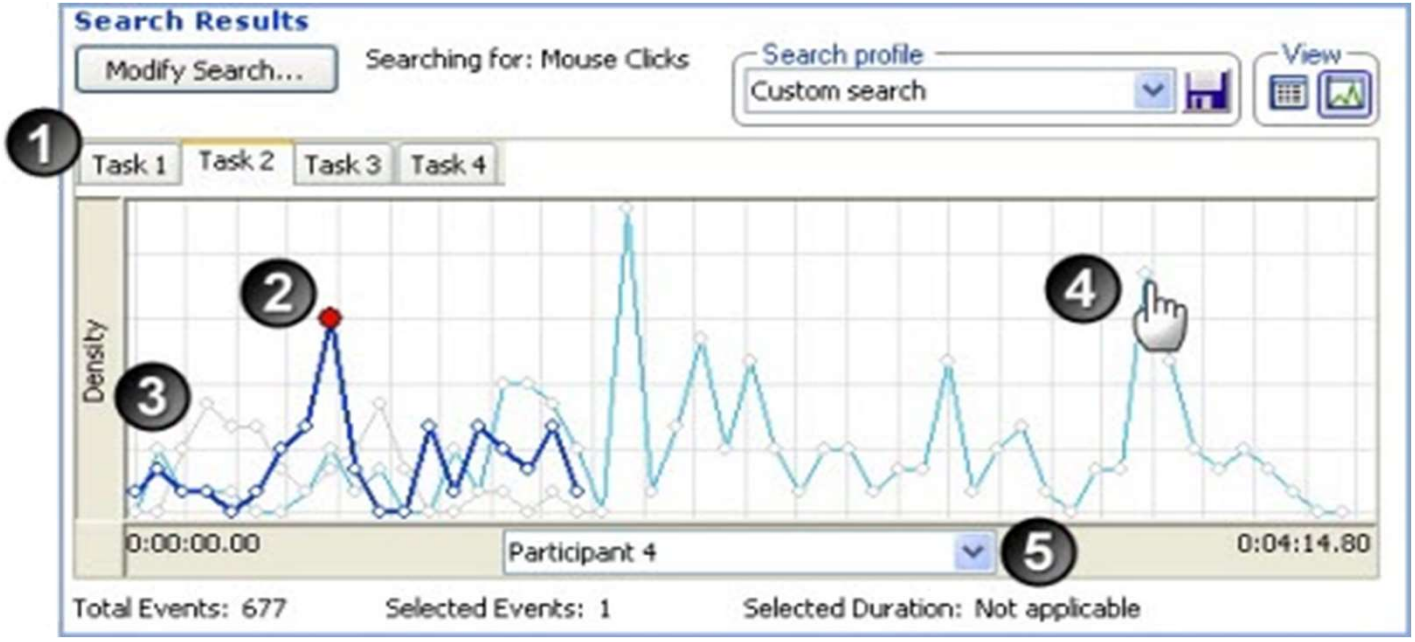
- Online or chat consultants, e-mail, and online suggestion boxes
  - Many users feel reassured if they know there is a human assistance available
  - On some network systems, the consultants can monitor the user's computer and see the same displays that the user sees

Evaluation During Active Use and Beyond (continued)

- Online suggestion box or e-mail trouble reporting
  - Electronic mail to the maintainers or designers
  - For some users, writing a letter may be seen as requiring too much effort
- Discussion groups, wikis and newsgroups
  - Permit postings of open messages and questions
  - Some are independent, e.g. America Online and Yahoo!
  - Topic list
  - Sometimes moderators
  - Social systems
  - Comments and suggestions should be encouraged

Evaluation During Active Use and Beyond (concluded)

- Example output of an automated evaluation tool from TechSmith's Morae
  - The item being measured is mouse clicks.
  - This shows the view for task 2 (selected in the tabbed bar). Obviously, the other 3 tasks could also be displayed. These are the values for participant 4.
  - The drop down list box would allow the evaluator to choose the mouse clicks for other participants.
  - Across the horizontal axis time is shown



Controlled Psychologically-oriented Experiments

- Scientific and engineering progress is often stimulated by improved techniques for precise measurement
- Rapid progress in the designs of interfaces will be stimulated as researchers and practitioners evolve suitable human-performance measures and techniques

Controlled Psychologically-oriented Experiments (continued)

- The outline of the scientific method as applied to human-computer interaction might comprise these tasks:
  - Deal with a practical problem and consider the theoretical framework
  - State a lucid and testable hypothesis
  - Identify a small number of independent variables that are to be manipulated
  - Carefully choose the dependent variables that will be measured
  - Judiciously select subjects and carefully or randomly assign subjects to groups
  - Control for biasing factors (non-representative sample of subjects or selection of tasks, inconsistent testing procedures)
  - Apply statistical methods to data analysis
  - Resolve the practical problem, refine the theory, and give advice to future researchers

Controlled Psychologically-oriented Experiments (concluded)

- Controlled experiments can help fine-tuning the human-computer interface of actively used systems
- Performance could be compared with the control group
- Dependent measures could include performance times, user-subjective satisfaction, error rates, and user retention over time