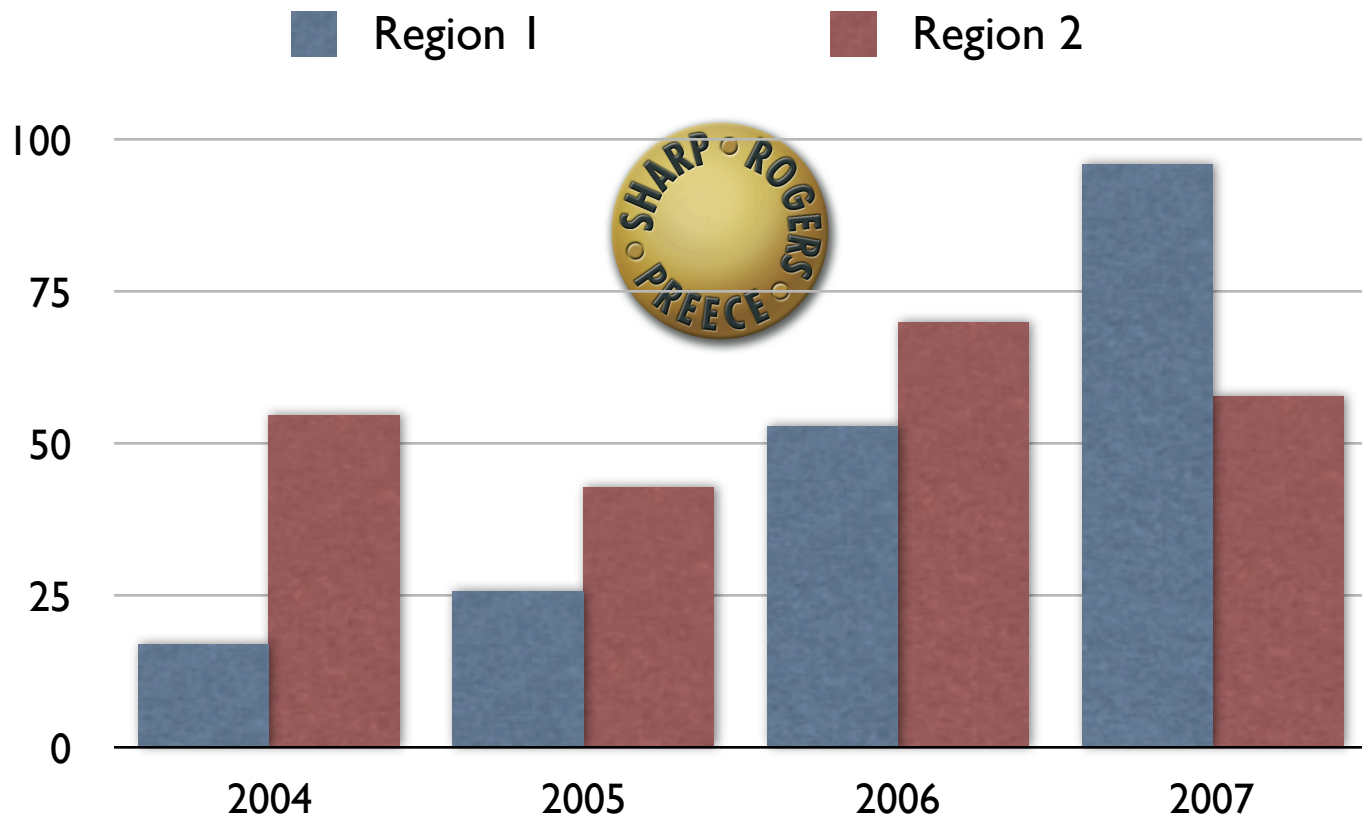


# User Interface Design

Dr. Oliver Obst

# An evaluation framework



# Formative Evaluation

- Find some users
  - Should be representative of the target user classes, based on user analysis
- Give each user some tasks
  - Should be representative of important tasks, based on task analysis
- Watch user do the tasks

# Two main types of evaluation

- **Formative evaluation** is done at different stages of development to check that the product meets users' needs.
- **Summative evaluation** assesses the quality of a finished product.

Our focus is on formative evaluation

# What to evaluate

- Iterative design & evaluation is a continuous process that examines:
  - Early ideas for conceptual model
  - Early prototypes of the new system
  - Later, more complete prototypes
- Designers need to check that they understand users' requirements.

# Why you need to evaluate:

- Bruce Tognazzini:

“Iterative design, with its repeating cycle of design and testing, is the only validated methodology in existence that will consistently produce successful results. If you don’t have user-testing as an integral part of your design process you are going to throw buckets of money down the drain.”

See <http://www.AskTog.com/> for discussion about design and evaluation.

# When to evaluate

- Throughout design
- From the first descriptions, sketches etc. of users' needs until the final product
- Design proceeds through iterative cycles of 'design-test-redesign'
- Evaluation is a key ingredient for a successful design.

# 1984 Olympic Message System

- Goal: Make it possible for athletes housed in Olympic Village dorms to send and receive recorded voice messages to and from each other, to coaches and staff outside the village, and to friends and relatives in all parts of the world
- Project initiated in late 1983
- In 1983-84 public voice message systems / answering machines were rare
- Operate for just 4 weeks; no time for a second try
- Failure would be broadcasted worldwide



# 1984 Olympic Message System

- Testing of prototypes started a few weeks into the project
- The early and constant availability of prototypes made extensive testing possible
  - On-site interviews/demos
  - Testing overseas with overseas users and telephone systems
  - Intensive usability tests with about 100 people
  - Local CS students were invited to come by and try to destroy the system
  - Field test at a pre-Olympic event
  - 2800 people at IBM Yorktown Heights and 1000 in Los Angeles were connected to the system and asked to use it
- And the system was a success

# 1984 Olympic Message System

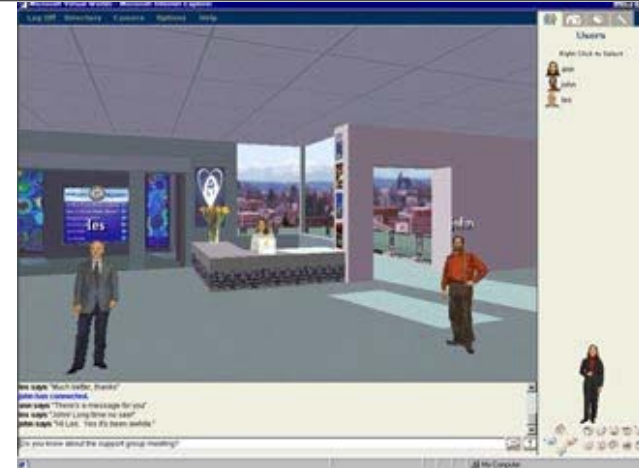
- The virtue of simplicity: many features dropped along the way, e.g.
  - Facility to verify that a message had been received
  - Facility to notify Olympians with waiting messages
  - Inserting new material in the middle of a message
  - Four alternatives in audio menus are too many
- Examples of problems encountered late in the project
  - Smog, heat, and desert bugs would create hardware problems
  - The user guide was translated to 10 additional languages
  - Names: Order of first, middle, and last name? Differing transcription rules

# Evaluating the 1984 OMS

- Early tests of printed scenarios & user guides
- Early simulations of telephone keypad
- An Olympian joined team to provide feedback
- Interviews & demos with Olympians outside US
- Overseas interface tests with friends and family
- Free coffee and donut tests
- Usability tests with 100 participants
- A 'try to destroy it' test
- Pre-Olympic field-test at an international event
- Reliability of the system with heavy traffic

# HutchWorld

- Virtual environment for cancer patients
- Reflect existing site (Fred Hutchinson Cancer Research Center in Seattle)
- Motivation: create a forum for patients and caregivers to:
  - talk about the disease
  - Find additional information resources
  - Communicate with family & friends
  - Stay in a sterile environment



# HutchWorld

- Many informal meetings with patients, caregivers & medical staff early in design
- Early prototype was informally tested on site
- Designers learned a lot, e.g.
  - language of designers & users was different
  - asynchronous communication was what the patients needed
- Redesigned to produce the portal version



# Usability Testing



- User tasks investigated:
  - how users' identity was represented
  - Communication
  - information searching
  - entertainment
- User satisfaction questionnaire
- Triangulation to get different perspectives

# Findings from the usability test



- The back button didn't always work
- Users didn't pay attention to navigation buttons
- Users expected all objects in the 3-D view to be clickable
- Users did not realise that there could be others in the 3-D world with whom to chat,
- Users tried to chat to the participant list
  - Other functions were triggered from this list
  - actual chat button was somewhere else

# Key Points

- Evaluation & design are closely integrated in user-centred design.
- Some of the same techniques are used in evaluation & requirements but they are used differently (e.g., interviews & questionnaires)
- Triangulation involves using a combination of techniques to gain different perspectives
- Dealing with constraints is an important skill for evaluators to develop

# The aims of this part of the lecture

- Explain key evaluation concepts & terms
- Describe the evaluation paradigms & techniques used in interaction design
- Discuss the conceptual, practical and ethical issues that must be considered when planning evaluations
- Introduce the DECIDE framework.

# What are user studies?

- looking at how people behave
  - either in their natural environments,
  - or in the laboratory,
- both
  - with old technologies and
  - with new/proposed technologies.

# Review / Overview of techniques

- observing users,
- asking users about their opinions,
- asking experts about their opinions,
- testing users' performance
- modelling users' task performance

# What is an evaluation paradigm?

- Any kind of evaluation is guided explicitly or implicitly by a set of beliefs, which are often under-pinned by theory.
- These beliefs and the methods associated with them are known as an '**evaluation paradigm**'

# Four evaluation paradigms

- 'quick and dirty'
- usability testing
- field studies
- predictive evaluation

# I. Quick and dirty

- **‘quick & dirty’ evaluation** describes the common practice in which designers informally
- get feedback from users or consultants to confirm that their ideas are in-line with users’ needs and are liked.
- Quick & dirty evaluations are done any time.
- The emphasis is on fast input to the design process rather than carefully documented findings.

# 2. Usability Testing

- Usability testing involves recording typical users' performance on typical tasks in controlled settings. Field observations may also be used
- As the users perform these tasks they are watched & recorded on video & their key presses are logged
- This data is used to calculate performance times, identify errors & help explain why the users did what they did
- User satisfaction questionnaires & interviews are used to elicit users' opinions

# Field Studies

- Field studies are done in natural settings
- The aim is to understand what users do naturally and how technology impacts them
- In product design field studies can be used to:
  - identify opportunities for new technology
  - determine design requirements
  - decide how best to introduce new technology
  - evaluate technology in use

# Predictive Evaluation

- Experts apply their knowledge of typical users, often guided by heuristics, to predict usability problems.
- Another approach involves theoretically based models.
- A key feature of predictive evaluation is that users need not be present
- Relatively quick & inexpensive

# DECIDE: a framework to guide evaluation

- **D**etermine the *goals* the evaluation addresses.
- **E**xplore the specific *questions* to be answered.
- **C**hoose the evaluation *approach* paradigm and *techniques* to answer the questions.
- **I**dentify the *practical issues*.
- **D**ecide how to deal with the *ethical issues*.
- **E**valuate, interpret and present the *data*.

# Determine the goals

- What are the high-level goals of the evaluation?
- Who wants it and why?
- The goals influence the approach used for the study
- Some examples of goals:
  - Identify the best metaphor on which to base the design
  - Check to ensure that the final interface is consistent
  - Investigate how technology affects working practices
  - Improve the usability of an existing product

# Explore the questions

- *All* evaluations need goals & questions to guide them, so time isn't wasted on ill-defined studies
- E.g., the goal of finding out why many customers prefer to purchase paper airline tickets rather than e-tickets can be broken down into sub-questions:
  - What are customers' attitudes to these new tickets?
  - Are they concerned about security?
  - Is the interface for obtaining them poor?
- What questions might you ask about the design of a cell phone?

# Choose the evaluation approach & techniques

- The evaluation *approach* strongly influences the *methods* used, and in turn, how data is collected, analysed and presented.
- E.g. field studies typically:
  - Involve observation and interviews.
  - Do not involve controlled tests in a laboratory.
  - Produce qualitative data.

# Identify practical issues

For example, how to:

- Select users
- Stay on budget
- Stay on schedule
- Find evaluators
- Select equipment

# Decide about ethical issues

- Develop an informed consent form
- Participants have a right to:
  - Know the goals of the study;
  - Know what will happen to the findings;
  - Privacy of personal information;
  - Leave when they wish;
  - Be treated politely.

# Evaluate, interpret & present data

- The approach and methods used influence how data is evaluated, interpreted and presented.
- The following need to be considered:
  - Reliability: can the study be replicated?
  - Validity: is it measuring what you expected?
  - Biases: is the process creating biases?
  - Scope: can the findings be generalised?
  - Ecological validity: is the environment influencing the findings? - e.g. Hawthorn effect.

# Pilot-studies

- A small trial run of the main study
- The aim is to make sure your plan is viable
- Pilot studies check:
  - that you can conduct the procedure
  - that interview scripts, questionnaires, experiments, etc. work appropriately
- It's worth doing several to iron out problems before doing the main study
- Ask colleagues if you can't spare real users

# Key Points

- An evaluation paradigm is an approach that is influenced by particular theories and philosophies.
- Five categories of techniques were identified: observing users, asking users, asking experts, user testing, modelling users.
- The DECIDE framework has six parts:
  - **D**etermine the overall goals
  - **E**xplore the questions that satisfy the goals
  - **C**hoose the paradigm and techniques
  - **I**dentify the practical issues
  - **D**ecide on the ethical issues
  - **E**valuate ways to analyse & present data
- Do a pilot study

# Confusion over Palm Beach County ballot

Although the Democrats are listed second in the column on the left, they are the third hole on the ballot.

Punching the second hole casts a vote for the Reform Party.

(REPUBLICAN) GEORGE W. BUSH - PRESIDENT DICK CHENEY - VICE PRESIDENT	3 →	(REFORM) PAT BUCHANAN - PRESIDENT EZOLA FOSTER - VICE PRESIDENT	← 4
(DEMOCRATIC) AL GORE - PRESIDENT JOE LIEBERMAN - VICE PRESIDENT	5 →	(SOCIALIST) DAVID McREYNOLDS - PRESIDENT MARY CAL HOLLIS - VICE PRESIDENT	← 6
(LIBERTARIAN) HARRY BROWNE - PRESIDENT ART OLIVIER - VICE PRESIDENT	7 →	(CONSTITUTION) HOWARD PHILLIPS - PRESIDENT J. CURTIS FRAZIER - VICE PRESIDENT	← 8
(GREEN) RALPH NADER - PRESIDENT WINDMA LaDUKE - VICE PRESIDENT	9 →	(WORKERS WORLD) MONICA MOOREHEAD - PRESIDENT GLORIA La RIVA - VICE PRESIDENT	← 10
(SOCIALIST WORKERS) JAMES HARRIS - PRESIDENT MARGARET TROWE - VICE PRESIDENT	11 →	WRITE-IN CANDIDATE To vote for a write in candidate, follow the directions on the long stub of your ballot card.	
(NATURAL LAW) JOHN HAGELIN - PRESIDENT NAT GOLDHABER - VICE PRESIDENT	13 →		

Sun-Sentinel graphic/Daniel Niblock