Chapter 12

Introducing Evaluation

Why, what, where and when to evaluate

Iterative design & evaluation is a continuous process that examines:

- What: a conceptual model, early prototypes of a new system and later, more complete prototypes.
- Where: in natural and laboratory settings.
- When: throughout design; finished products can be evaluated to collect information to inform new products.
- Why: to check users' requirements, that users can use the product, and that they like the product.

Bruce Tognazzini tells you why you need to evaluate

"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 AskTog.com for topical discussions about design and evaluation.

Types of evaluation

- Controlled settings involving users
  - usability testing & experiments in laboratories and living labs.
- Natural settings involving users
  - field studies to see how the product is used in the real world.
- Any settings not involving users
  - consultants critique; to predict, analyze & model aspects of the interface analytics.

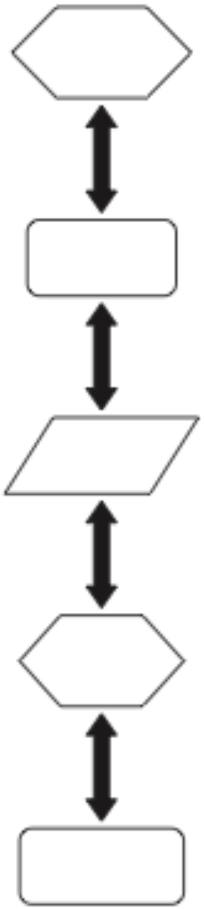
Usability lab

http://iat.ubalt.edu/usability_lab/

Living labs

- People's use of technology in their everyday lives can be evaluated in living labs.
- Such evaluations are too difficult to do in a usability lab.
- Eg the Aware Home was embedded with a complex network of sensors and audio/video recording devices (Abowd et al., 2000).

Usability testing & field studies can compliment



Field study to evaluate initial design ideas and get early feedback

Make some design changes

Usability test to check specific design features

Field study to see what happens when used in natural environment

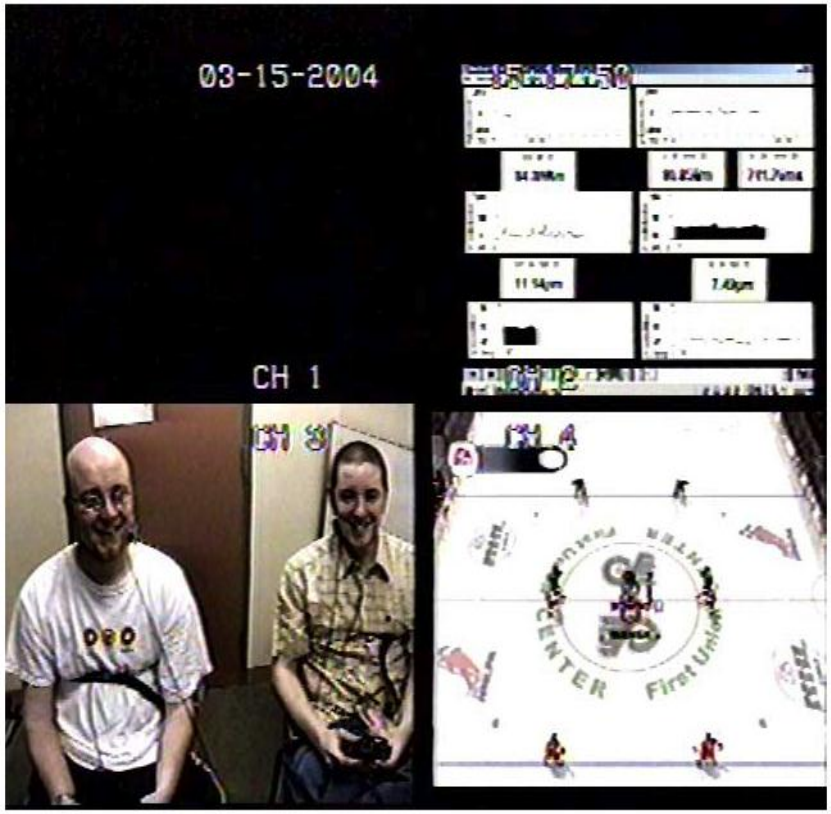
Make some final design changes

Evaluation case study

- Experiment to investigate a computer game

Challenge & engagement in a collaborative immersive game

- Physiological measures were used.
- Players were more engaged when playing against another person than when playing against a computer.



What does this data tell you?

	Playing against computer		Playing against friend	
	Mean	St. Dev.	Mean	St. Dev.
Boring	2.3	0.949	1.7	0.949
Challenging	3.6	1.08	3.9	0.994
Easy	2.7	0.823	2.5	0.850
Engaging	3.8	0.422	4.3	0.675
Exciting	3.5	0.527	4.1	0.568
Frustrating	2.8	1.14	2.5	0.850
Fun	3.9	0.738	4.6	0.699

Source: Mandryk and Inkpen (2004).

Evaluation methods

Method	Controlled settings	Natural settings	Without users
Observing	x	x	
Asking users	x	x	
Asking experts		x	x
Testing	x		
Modeling			x

The language of evaluation

- Analytics
- Analytical evaluation
- Controlled experiment
- Expert review or crit
- Field study
- Formative evaluation
- Heuristic evaluation
- In the wild evaluation
- Living laboratory
- Predictive evaluation
- Summative evaluation
- Usability laboratory
- User studies
- Usability testing
- Users or participants

Key points

- Evaluation & design are closely integrated in user-centered design.
- Some of the same techniques are used in evaluation as for establishing requirements but they are used differently (e.g. observation interviews & questionnaires).
- Three types of evaluation: laboratory based with users, in the field with users, studies that do not involve users
- The main methods are: observing, asking users, asking experts, user testing, inspection, and modeling users' task performance, analytics.
- Dealing with constraints is an important skill for evaluators to develop.