Chapter 12

Introducing Evaluation

Main aims of this chapter…

- Explain the key concepts and terms used in evaluation.
- Introduce a range of different types of evaluation methods.
- Show how different evaluation methods are used for different purposes at different stages of the design process and in different contexts of use.
- Show how evaluators mix and modify methods to meet the demands of novel systems.
- Discuss some of the practical challenges that evaluators have to consider when doing evaluation.
- Illustrate through short case studies how methods discussed in more depth in Chapters 7 and 8 are used in evaluation and describe some methods that are specific to evaluation.

Why, what, where and when to evaluate

Iterative design & evaluation is a continuous process that examines:

- Why: to check users' requirements and that users can use the product and they like it.
- 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.

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.

The language of evaluation

- Analytics
- Analytical evaluation
- Controlled experiment
- Expert review
- 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

Types of evaluation

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

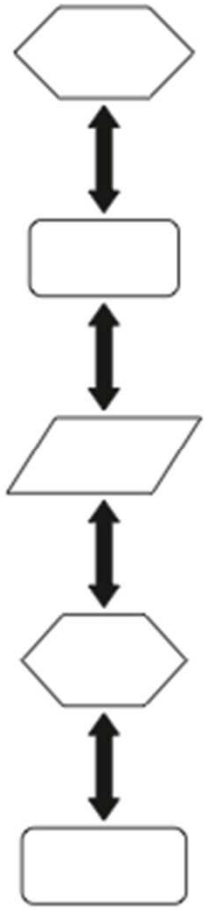
Characteristics of approaches

	Usability testing	**Field studies**	**Analytical**
Users	do task	natural	not involved
Location	controlled	natural	anywhere
When	prototype	early	prototype
Data	quantitative	qualitative	problems
Feed back	measures & errors	descriptions	problems
Type	applied	naturalistic	expert

Evaluation methods

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

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

Usability lab

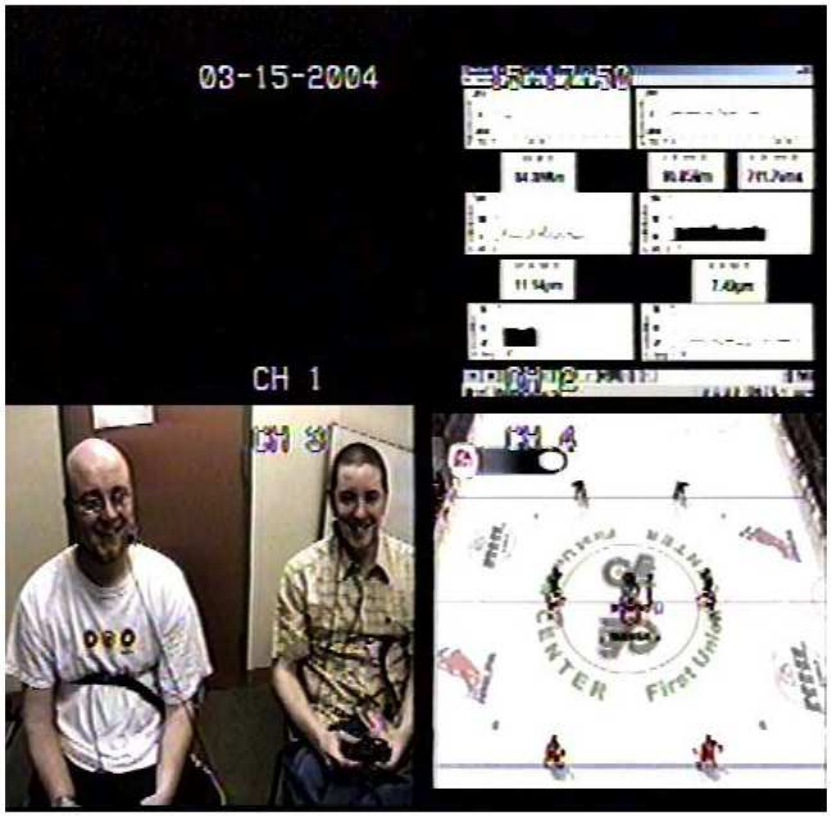
http://iat.ubalt.edu/usability_lab/

Evaluation case studies

- Experiment to investigate a computer game
- In the wild field study of skiers
- Crowdsourcing

Challenge & engagement in a collaborative immersive game

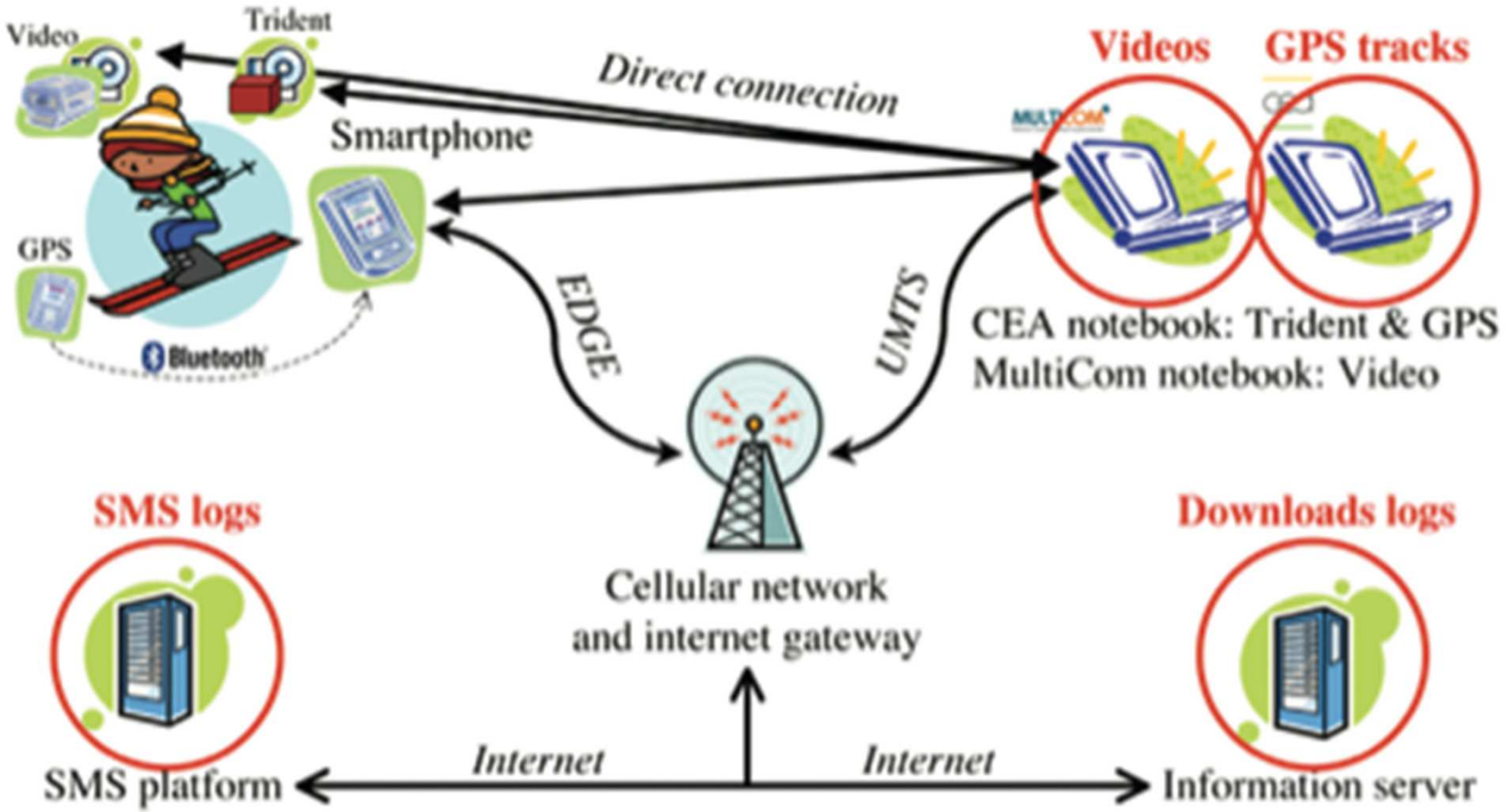
- Physiological measures were used.
- Players were more engaged when playing against another person than when playing against a computer.
- What precautionary measures did the evaluators take?



Why study skiers in the wild?

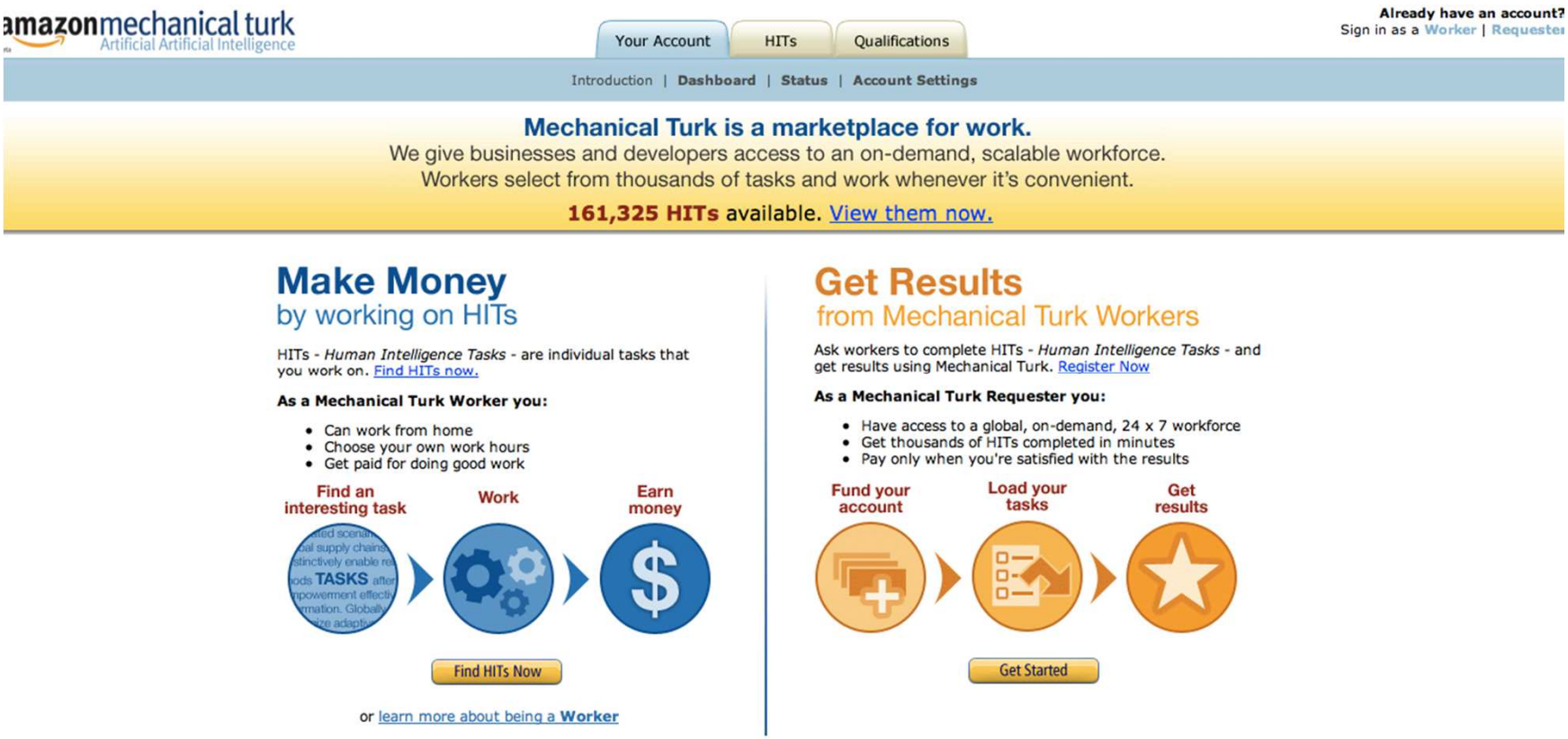
Jambon et al. (2009) User experience in the wild. In: Proceedings of CHI ’09, ACM Press, New York, p. 4070-4071.

e-skiing system components



Jambon et al. (2009) User experience in the wild. In: Proceedings of CHI '09, ACM Press, New York, p. 4072.

Crowdsourcing - when might you use it?



Summary

- 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.