BINUS UNIVERSITY

People
Innovation
Excellence

Course : COMP6176 / Human - Computer Interaction

Year : 2019

BASIC EVALUATION

SESSION 11

LEARNING OUTCOMES

- LO 5: Evaluate the user interfaces of interactive software

People
Innovation
Excellence

BINUS UNIVERSITY

People
Innovation
Excellence

OUTLINE

- Introduction
- The Why, What, Where, and When of Evaluation
- Types of Evaluation
- Evaluation Case Studies

BINUS UNIVERSITY

People
Innovation
Excellence

INTRODUCTION

- **Evaluation** is integral to the design process.
- Evaluator collect information about users' or potential users' experiences when interacting with a prototype, a computer system, a component of a computer system, an application or a design artifact such as a screen sketch.
- Evaluation focuses on both the usability of the system (e.g. how easy it is to learn and to use) and on the user's experience when interacting with the system (e.g how satisfying, enjoyable, or motivating the interaction is)

BINUS UNIVERSITY

People
Innovation
Excellence

INTRODUCTION

- Users expect much more than just a usable system, they look for a pleasing and engaging experience, that why evaluation is important to carry out.
- What to evaluate ranges from low-tech prototypes to complete system, a particular screen function to the whole workflow, and from aesthetic design to safety features.
- Where evaluation takes place depends on what is being evaluated. Some characteristics such as web accessibility generally evaluated in a laboratory, or user experience aspects can be evaluate effectively in natural settings which often refer as in the wild studies.

BINUS UNIVERSITY

People
Innovation
Excellence

INTRODUCTION

- When evaluate or at what stage in the product lifecycle evaluation takes place depends on the type of product.
- When evaluations are done during design to check that a product continues to meet user's need they are known as formative evaluations.
- Evaluations that are done to assess the success of a finished product are known as summative evaluations.
- If the products is being upgraded then the evaluation may not focus on establishing a set of requirements, but may evaluate the existing product to ascertain what needs improving.

BINUS UNIVERSITY

People
Innovation
Excellence

The Why, What, Where and When of Evaluation

Why Evaluate?

- User experience involves all aspects of the user's interaction with the product.
- Nowadays users expect much more than just a usable system—they also look for a pleasing and engaging experience from more products.
- Simplicity and elegance are valued so that the product is a joy to own and use.

BINUS UNIVERSITY

People
Innovation
Excellence

The Why, What, Where and When of Evaluation

Why Evaluate?

- From a business and marketing perspective, well-designed products sell. Hence, there are good reasons for companies to invest in evaluating the design of products.
- Designers can focus on real problems and the needs of different user groups and make informed decisions about the design

BINUS UNIVERSITY

People
Innovation
Excellence

The Why, What, Where and When of Evaluation

What to Evaluate?

- What to evaluate ranges from low-tech prototypes to complete systems, from a particular screen function to the whole workflow, and from aesthetic design to safety features.
- Developers of a new web browser may want to know whether users find items faster with their product.

BINUS UNIVERSITY

The Why, What, Where and When of Evaluation

What to Evaluate?

- Example: Game app developers will want to know how engaging and fun their games are compared with those of their competitors and how long users will play them.

BINUS UNIVERSITY

People
Innovation
Excellence

The Why, What, Where and When of Evaluation

Where to Evaluate?

- Where evaluation takes place depends on what is being evaluated.
- Some characteristics, such as web accessibility, are generally evaluated in a lab because it provides the control necessary to investigate systematically whether all of the requirements are met.

The Why, What, Where and When of Evaluation

Where to Evaluate?

- Example: Remote studies of online behavior, such as social networking, can be conducted to evaluate natural interactions of participants in the context of their interaction, for example, in their own homes or place of work.

BINUS UNIVERSITY

People
Innovation
Excellence

The Why, What, Where and When of Evaluation

When to Evaluate?

- The stage in the product lifecycle when evaluation takes place depends on the type of product and the development process being followed.
- Example : the product being developed could be a new concept, or it could be an upgrade to an existing product.
- If the product is new, then considerable time is usually invested in market research and discovering user requirements.

BINUS UNIVERSITY

People
Innovation
Excellence

The Why, What, Where and When of Evaluation

When to Evaluate?

- When evaluations are conducted during design to check that a product continues to meet users' needs, they are known as **formative evaluations.**
- **Formative evaluations** cover a broad range of design processes, from the development of early sketches and prototypes through to tweaking and perfecting a nearly finished design.

BINUS UNIVERSITY

People
Innovation
Excellence

The Why, What, Where and When of Evaluation

When to Evaluate?

- Evaluations that are carried out to assess the success of a finished product are known as **summative evaluations**.
- If the product is being upgraded, then the evaluation may not focus on discovering new requirements but may instead evaluate the existing product to ascertain what needs improving.

BINUS UNIVERSITY

People
Innovation
Excellence

TYPES OF EVALUATION

- These are three classification of evaluation depending on the setting, user involvement and level of control :
  1. Controlled settings directly involving users (ex : laboratory and living labs)
  2. Natural settings involving users (ex : on line communities and public places)
  3. Any setting not directly involving users (ex : consultants, researchers critique)

Controlled Setting Involving Users

- Controlled settings enable evaluator to control what users do, when they do it, and for how long. Also enables them to reduce outside influences and distraction, such as colleagues talking.

BINUS UNIVERSITY

People
Innovation
Excellence

TYPES OF EVALUATION

- This approach used to **evaluate software applications running on PC and other technologies** where participants can be seated in front of them to perform a set of tasks.
- The main method are **usability testing and experiments.**
- **Usability testing** involves collecting data using a combination of method - i.e experiments, observation, interviews, questionnaires - in a controlled setting.
  - The Primary Goals is to determine whether an interface is usable by the intended user population to carry out the tasks for which it was designed.

TYPES OF EVALUATION

Natural Settings Involving Users

- The main method is field studies. The aim of field studies is to evaluate people in their natural settings.
- Used primarily to :
  - Help identify opportunities for new technology
  - Establish the requirements for a new design
  - Facilitate the introduction of technology or inform deployment of existing technology in new contexts.
- Methods are typically used in field study are observations, interviews, and logging.
- The data takes the form of events and conversations that are recorded by the researchers as notes or by audio or video recording, or by participants as diaries and notes

BINUS UNIVERSITY

People
Innovation
Excellence

TYPES OF EVALUATION

Natural Settings Involving Users

- **A goal** is to be unobtrusive and not to affect what people do during the evaluation.
- Field study can also be virtual, where observations take place in multiuser games such as World of Warcraft, online communities, chat room.
  - A goal of this kind of field study is to examine the kinds of social process that occur in them, such as collaboration, confrontation, and cooperation.

BINUS UNIVERSITY

People
Innovation
Excellence

TYPES OF EVALUATION

Any settings Not involving Users

- This Evaluations are conducted in settings where the researcher has to imagine or model how an interface is likely to be used.
- The method will be used are :
  - **Inspection Method** to predict user behavior and to identify usability problems, based on knowledge of usability, users' behavior, the contexts in which the system will be used, and the kinds of activities that users undertake.
  - **Examples include heuristic evaluation** that applies knowledge of typical users guided by rules of thumb
  - **Walkthroughs** that involve stepping through a scenario or answering a set of question for a detailed prototype.

  Other techniques include **analytics and models**

BINUS UNIVERSITY

People
Innovation
Excellence

TYPES OF EVALUATION

Any settings Not involving Users

- Other techniques include **analytics and models**.
  - **Analytics** is a technique for logging data either at a customer's site or remotely. Web analytics is the measurement, collection, analysis, and reporting of Internet data in order to understand and optimize web usage (Arikan,2008).
  - **Models** have been used primarily for comparing the efficacy of different interfaces for the same applications, for example **Fitts'Law** to predict the time it takes to reach a target using a pointing device.
- The three broad categories identified above provide a general framework to guide the selection of evaluation methods.

BINUS UNIVERSITY

People
Innovation
Excellence

TYPES OF EVALUATION

- Often combinations of methods are used across the categories to obtain a richer understanding.
- There are obviously pros and cons between using a controlled and uncontrolled setting.
- **A benefit of uncontrolled settings** is that unexpected data can be obtained that provides quite different insights into people's perceptions and their experiences of using, interacting , or communicating through the new technologies in the context of their everyday and working lives.
- **A benefit of controlled settings** include being able to test hypotheses about specific features of the interface, where the results can be generalized to the wider population.

BINUS UNIVERSITY

People
Innovation
Excellence

EVALUATION CASE STUDIES

- This case study describes an experiment that tested whether it was more exciting playing against a computer versus playing against a friend for a collaborative computer game (MandryK and Inkpen,2004).
- Case Study : An experiment investigating a computer game.

  For games to be successful they must engage and challenge users. Ways of evaluating this aspect of the user experience are therefore needed and , in this case study (Mandryk and Inkpen,2004), physiological responses were used to evaluate users' experiences when playing against a friend and when playing alone against the computer. The researchers conjectured that physiological indicators could be an effective way of measuring a player's experience. Specifically they designed an experiment to evaluate the participants' experience of playing an online ice-hockey game. *(For complete study case see textbook Sub bab 12.4 page 443-444)*

EVALUATION CASE STUDIES

- Figure 11.01 shows the set-up for recording data while the participants were playing the game.

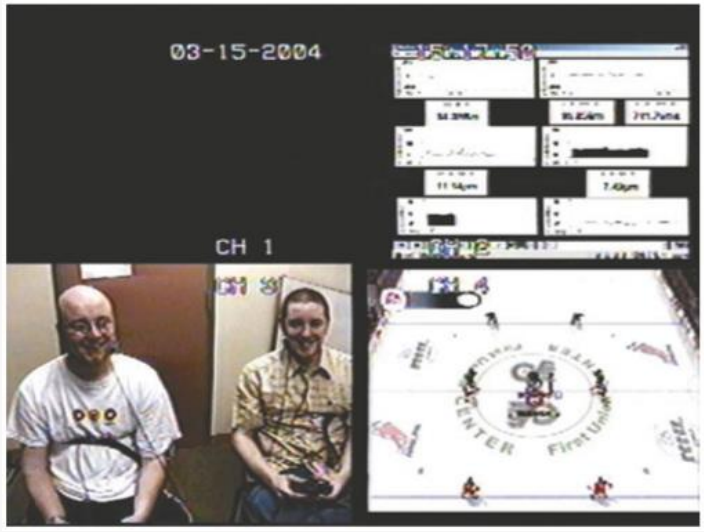


Figure 11.01 The display show the physiological data (top right), two participants and a screen of the game they played

- Result from the user satisfaction questionnaire revealed that the mean ratings on a 1 – 5 scale for each item indicated that playing against a friend was the favored experience (Table 11.1) .

People
Innovation
Excellence

BINUS UNIVERSITY

People
Innovation
Excellence

EVALUATION CASE STUDIES

- Table 11.01 Mean subjective ratings given on a user satisfaction questionnaire using a five-point scale.

	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

- Based on table 11.01, Identifying strongly with an experience state is indicated by a higher mean. The standard deviation indicates the spread of the results around the mean. Low values indicate little variation in participants' responses, high values indicate more variation.

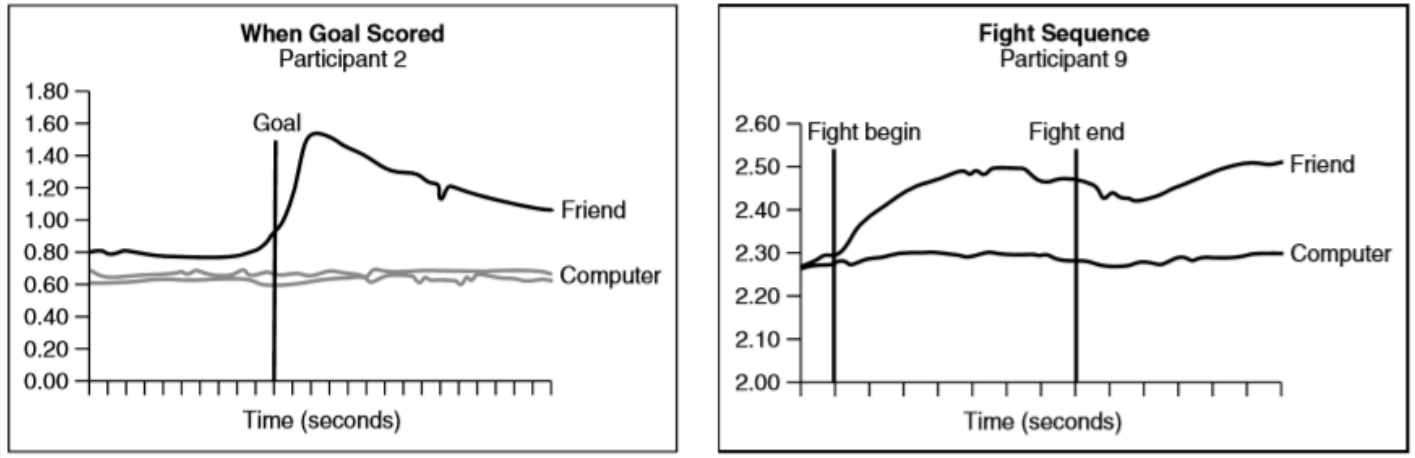
BINUS UNIVERSITY

People
Innovation
Excellence

EVALUATION CASE STUDIES

- The physiological recordings were also compared across participants and in general indicated the same trend. (The comparison for two participants has shown in Fig. 11.02)
- In Figure 11.02 (a) the result of a participant's skin response when scoring a goal against a friend versus against the computer.
- In Figure 11.02 (b) another participant's response when engaging in a hockey fight against a friend versus against the computer.

EVALUATION CASE STUDIES



- (a) (b)

Figure 11.02 (a) and (b) comparison for two participants

- From the evaluation result, this indicated that the physiological data gathering and analysis methods were effective for evaluating levels of challenge and engagement.

BINUS UNIVERSITY

People
Innovation
Excellence

REFERENCES

- Interaction Design 5th Edition 2019, Chapter 14
- https://medium.com/@aaronlee_83694/interaction-design-chapter-13-introducing-evaluation-2-21-19-a46a52866900
- https://www.cs.bham.ac.uk/internal/courses/hci/lectures/HCI%20-%20Evaluation.pdf