8 : Evaluating Designs

IT4106 – User Experience Design (UXD)

Level II - Semester 4

Overview

- The goal of this topic is to introduce the importance of evaluation, what needs to be evaluated, where evaluation should take place, and when in the product lifecycle evaluation is needed.

Intended Learning Outcomes

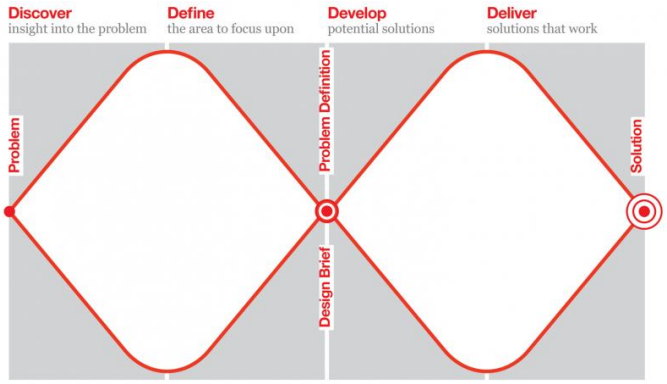
- At the end of this lesson, you will be able to;
  - Explain the key concepts and terms used in evaluation
  - Introduce a range of different types of evaluation methods
  - Show how different evaluation methods are used for different purposes at different stages of the design process and in different contexts of use
  - Show how evaluation methods are mixed and modified to meet the demands of evaluating novel systems

List of sub topics

- 1.1 Introduction to Design Evaluation
  - 1.1.1 Evaluating in Controlled Settings
  - 1.1.2 Evaluating in Natural Settings
  - 1.1.3 Evaluating without users
- 1.2 Usability Testing
- 1.3 Conducting User Experience Experiments
- 1.4 Conducting Field Trials
- 1.5 Other Evaluation methods
  - 1.5.1 Heuristic Evaluation
  - 1.5.2 Walk-Throughs
    - 1.5.2.1 Cognitive Walk-Throughs
    - 1.5.2.2 Pluralistic Walk-Throughs
    - 1.5.2.3 Web Analytics
    - 1.5.2.4 A/B Testing

Introduction to Design Evaluation

- Last phase of double diamond
- Evaluation is integral to the design process.
  - collecting and analyzing data about users' experiences when interacting with a design artifact
- Goal of evaluation is to improve the artifact's design.
- Evaluation focuses on both the usability of the system and on the users' experiences when interacting with it.



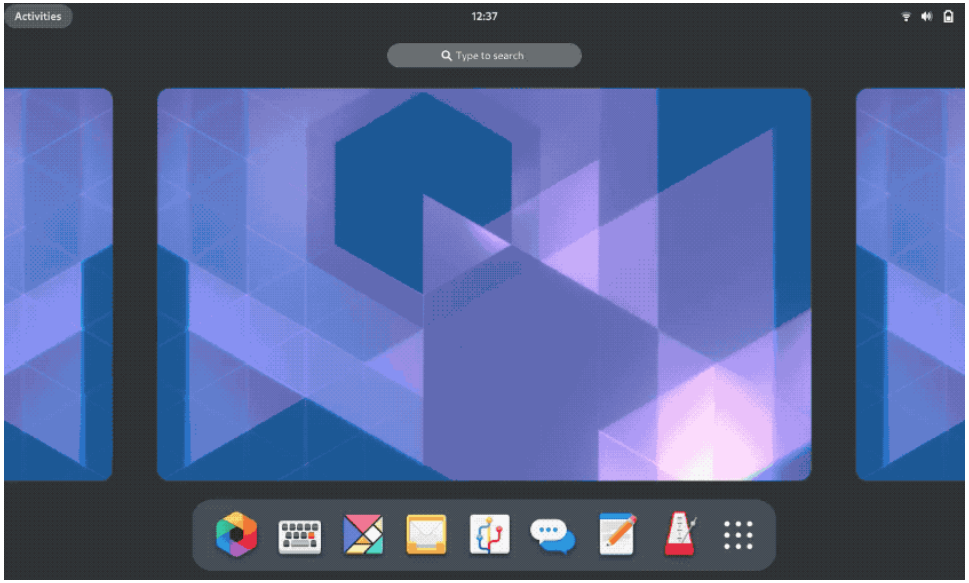
Introduction to Design Evaluation

- There are many different evaluation methods
  - depends on the goals of the evaluation
- Usually involve observing participants
- Methods that do not involve users such as modeling users' behavior and analytics
- There are different levels of control
  - In the wild
  - Lab environment

Illustration by slidesgo on freepik

Why Evaluate?

- User experience involves all aspects of the user's interaction with the product
  - users expect much more than just a usable system
- Easy to sell well-designed products

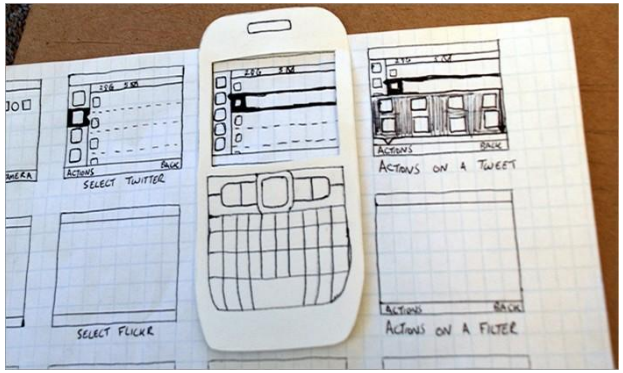


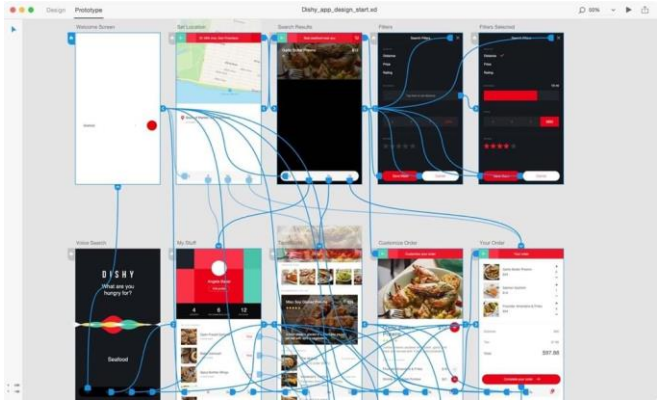
Ubuntu launcher changes then and now

"Simply improving customer journeys has the potential to increase customer satisfaction by 20% but also to lift revenue by up to 15% while lowering the cost of serving customers by as much as 20%." – McKinsey

What to Evaluate

- Ranges from low-tech prototypes to complete systems.





- Different types of evaluations will be needed depending on the type of product

Where to Evaluate

- Where evaluation takes place depends on what is being evaluated
  - Lab environment : characteristics, such as web accessibility
  - In a natural setting – AKA in-the-wild studies. : User experience aspects
  - Remote studies : online behavior, such as social networking

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.
  - E.g. If the product is new, then considerable time is usually invested in market research and discovering user requirements. Then create prototypes and evaluate afterwards
- **Formative evaluations:** When evaluations are conducted during design to check that a product continues to meet users' needs
- **Summative evaluations**: When Evaluations that are carried out to assess the success of a finished product

Usability Testing

- Traditionally been tested in controlled laboratory settings.
  - Emphasizes how usable a product is.
- Enables designers to control what users do and allows them to control the environmental and social influences that might impact the user's performance.
- Test whether the product being developed is usable by the intended user in order to achieve the tasks for which it was designed and whether users are satisfied with their experience.

Types of Evaluation - depending on the setting, user involvement, and level of control

- Controlled settings directly involving users:
  - Users' activities are controlled to test hypotheses and measure or observe certain behaviors.
  - The main methods are usability testing and experiments.
- Natural settings involving users:
  - There is little or no control of users' activities to determine how the product would be used in the real world.
  - The main method used is field studies.
- Any settings not directly involving users:
  - Consultants and researchers' critique, predict, and model aspects of the interface to identify the most obvious usability problems.
  - The range of methods includes inspections, heuristics, walk-throughs, models, and analytics.

Evaluating in Controlled Settings

- Experiments and user tests are designed to control what users do, when they do it, and for how long.
- Evaluating user interfaces involves collecting data using a combination of methods in a controlled setting.
- Determine whether an interface is usable by the intended user population to carry out the tasks for which it was designed.

Evaluating in Natural Settings

- Evaluate products with users in their natural settings.
- The data takes the form of events and conversations that are recorded by the researchers as notes, or through audio or video recording, or by the participants as diaries and notes.
- The goal is to be unobtrusive and not to affect what people do during the evaluation.

Crowdsourcing

- The Internet makes it possible to gain access to hundreds of thousands of participants who will perform tasks or provide feedback on a design or experimental task quickly and almost immediatelv.

Evaluating without users

- Evaluations that take place without involving users are conducted in settings where the researcher has to imagine or model how an interface is likely to be used.
- Data Collected via;
  - Knowledge codified in heuristics
  - Data collected remotely
  - Models that predict users' performance
- Popular methods
  - Heuristic evaluation and walkthroughs
  - Analytics and A/B testing
  - Predictive modeling

Heuristic Evaluation

- Guided by a set of usability principles known as heuristics, evaluate whether user-interface elements, such as dialog boxes, menus, navigation structure, online help, and so on, conform to tried-and-tested principles.
- These heuristics closely resemble high-level design principles.

10 Usability Heuristics for User Interface Design (nngroup.com)

Heuristic Evaluation Follows - Jakob Nielsen's 10 general principles for interaction design

- Visibility of System Status
- Match Between System and the Real World
- User Control and Freedom
- Consistency and Standards
- Error Prevention
- Recognition Rather Than Recall
- Flexibility and Efficiency of Use
- Aesthetic and Minimalist Design
- Help Users Recognize, Diagnose, and Recover from Errors
- Help and Documentation

Walk-Throughs

- Offer an alternative approach to heuristic evaluation for predicting users' problems without doing user testing.
- Involve walking through a task with the product and noting problematic usability features.
- Most walk-through methods do not involve users.

Cognitive Walk-Throughs

- Cognitive walk-throughs involve simulating how users go about problem-solving at each step in a human-computer interaction.
- Takes a cognitive perspective in which the focus is on evaluating designs for ease of learning
- Focus that is motivated by observations that users learn by exploration.

Cognitive Walk-Throughs –An example

1. Withdraw 6000 LKR from their savings account
   I. Insert a debit card
   II. Enter PIN
   III. Tap 6000 LKR of the screen
2. Check the balance in the account
3. And deposit a check.

Fours questions to ask

1. Will the user try and achieve the right outcome?
2. Will the user notice that the correct action is available to them?
3. Will the user associate the correct action with the outcome they expect to achieve?
4. If the correct action is performed; will the user see that progress is being made towards their intended outcome?

Pluralistic Walk-Throughs

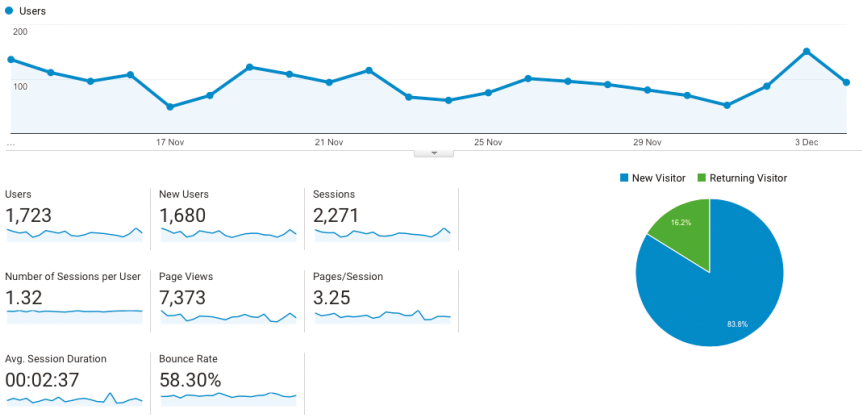
- Users, developers, and usability researchers work together to step through a task scenario.
- Discuss usability issues associated with dialog elements involved in the scenario steps.
- Each person is asked to assume the role of a typical user.
- Scenarios of use, consisting of a few prototype screens, are given to each person who writes down the sequence of actions that they would take to move from one screen to another, without conferring with each other.
- The benefits of pluralistic walk-throughs include a strong focus on users' tasks at a detailed level, that is, looking at the steps taken.

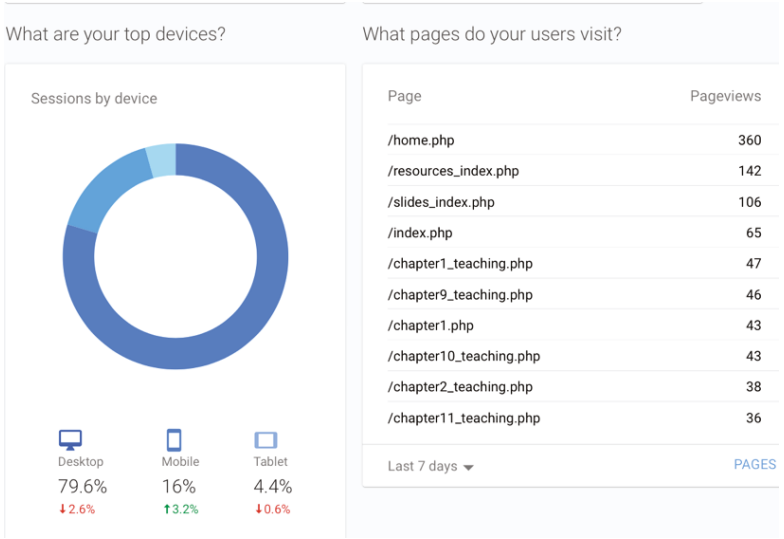
Analytics and A/B Testing

- Users' actions can be recorded by software
  - key presses
  - mouse or other pointing device movements
  - time spent searching a web page
  - looking at help systems
  - task flow through software modules

Web Analytics

- A form of interaction logging that was specifically created to analyze users' activity on websites
  - If the website is used by thousands of users and a small number of users do not return, this loss of users may not be noticed by the web designers and web owners unless they track users' activities.
  - Ex: Google Analytics





A/B Testing

- Evaluate a website, part of a website, an application, or an app running on a mobile device
- Carrying out a large-scale experiment to evaluate how two groups of users perform using two different designs, one of which acts as the control and the other as the experimental condition.
- It is basically a controlled experiment but one that often involves hundreds or thousands of participants.

A/B Testing – Example - Ron Kohavi and Roger Longbotham (2015)

- Comparing early versions of Microsoft Office 2007 home page
- To test the effectiveness of a new and more modern-looking home page
  - To increase the number of download clicks
- Downloads decreased by 64% due to design issue
  - The words in the new design were “Buy now” with a $149.95 price, whereas the old design said, “Try 2007 for free” and “Buy now”.

A

B

A/B testing is a fundamental and critical Web services… consistent use of A/B testing could save the company millions of dollars - The Office Online team