Digital Interaction Design IS-104

EVALUATION

MOHAMMAD TAFIQUR RAHMAN

OCTOBER 30, 2019

UiA University of Agder
Norway

Recap from last lecture

UI design principles:

- Organize: a clear and consistent conceptual structure
- Economize: do the most with least amount of cues
- Communicate: match the presentation to the capabilities of the user

UI Stack:

- Ideal State: the product is working properly
- Error State: when things go wrong
- Partial State: no longer empty and sparsely populated
- Loading State: represent situation when fetching data
- Blank State: the first impression to the user

Envisionment:

- Brings designs to life for both designers and the people who will use the new designs.
- The point is to explore ideas, not to build an entire parallel system or product.
- Different methods for "envisionment": Storyboards, Navigation Maps, Wireframes and Prototypes.
- Prototypes can be lo-fi or hi-fi basing on:
  - the stage of the project and
  - the context for the use of the prototype.

University of Agder
Norway

Goals for today
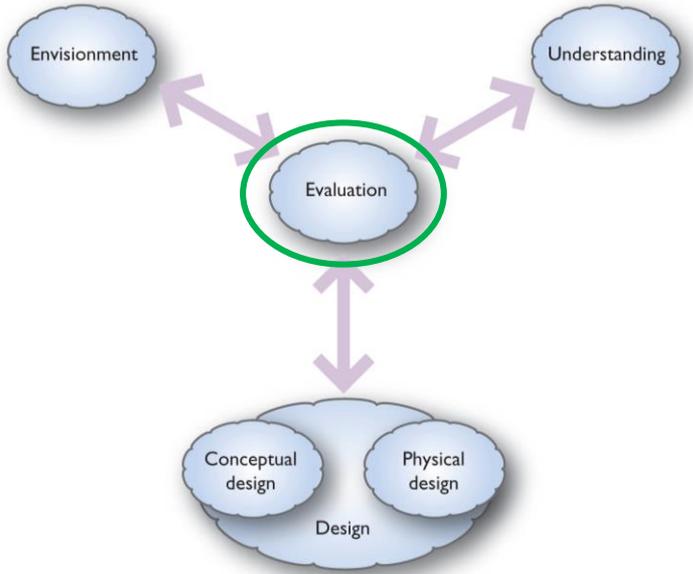
Design Evaluation

- Why centrally positioned?

- What are evaluation criteria?

- Methods for evaluation:
  - Expert-based
  - Participant-based
  - Data Analytics

University of Agder
Norway

Evaluation – the core of UX designing activity

Evaluation:

- Central to whole design
  - It reviews, tries out or tests
    - design ideas, software, products or services
    - to discover whether it meets some criteria

- UX designers are concerned
  - Not only with surface features: icons, elements, colours, layouts
  - But also the system is fit for its purpose: enjoyable, engaging, problem solving
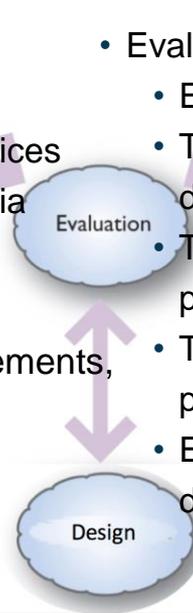
- Evaluation is not just for finished systems:
  - Early ideas can initially be evaluated for envisioning
  - The envisioned idea can further be evaluated for designing
  - The evaluated design can procced to product prototyping
  - The prototype can finally be evaluated for system (or product) development
  - Evaluation of finished systems can be performed before delivery

Envisionment

Understanding

Evaluation
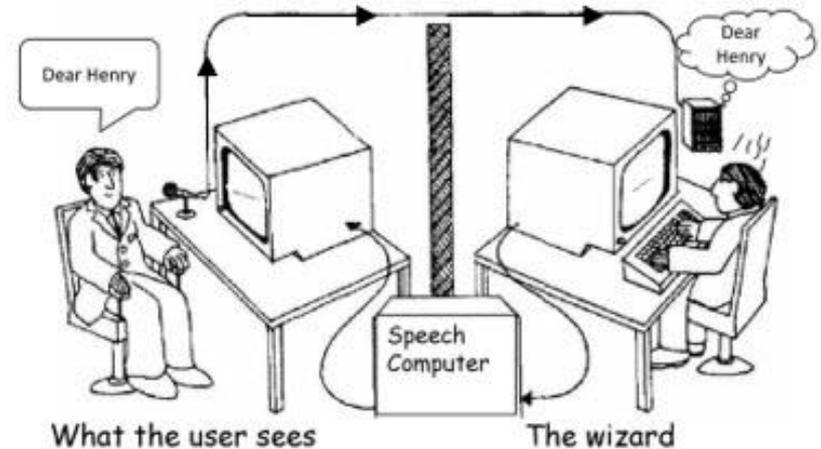
Design

Evaluation – good/bad idea

- Evaluating our own thinking is a habit that few of us practice.
- In user-centred approach, designs are evaluated right from the earliest idea.
- Evaluation is an integral part of an evolutionary design process.
- We use evaluations to rework parts of the design or to decide between options (for example, between voice input or touchscreen interaction for elderly with dementia).

Evaluation – when?

- Evaluation is dependent on the form of envisionment - used to represent the system.
- It is only possible to evaluate what is present in the available system version.
- But it is possible to imitate convincingly the system functionality without developing it.
- Evaluation occurs throughout the interaction design process.
- For example, Wizard of Oz evaluations, and other lo-fi evaluations.

Wizard of Oz testing – The listening type writer IBM 1984

Dear Henry

Dear Henry

Speech Computer

What the user sees

The wizard

Evaluating Criteria

- Evaluation entails assessing a design against specific criteria
- The criteria can be derived from:
  - Formal design principles (e.g. Benyon's (chap 5) or Nielsen´s)
  - Standards (e.g. ISO 9241 on ergonomics of human system interaction).
  - Customer defined requirements and recommendations (e.g., a maximum time to complete a set of operations, or DIFI´s assessment criteria for public digital services).
  - In-house defined guidelines.
- The criteria used can differ depending on the **aims** of the system and the **maturity** of the design.

Types of Evaluation

- Three main types of Evaluation:

  - Expert-based evaluation
    - if usability experts, or interaction designers evaluate some envisioned version of a design.
    - also called inspection methods.

  - Participant-based evaluation
    - if people from the anticipated user groups evaluate some envisioned version of a design.
    - also called user testing methods.

  - Data analytics
    - gather data on system performance when it is deployed.

Expert-based Evaluation

- Relatively quick, effective and cheap.
- No substitute for asking real users to participate in evaluation but can be useful particularly early in the design process.
- Experts can pick up problems before a lot of effort is made towards a specific direction.
- They utilize their experiences to identify affecting factors
- Two main approaches:
  - Heuristic evaluation: is performed against a list of principles or heuristics
  - Cognitive walkthrough: checks detailed design and logic of steps in an interaction

Heuristics Evaluation

- A person trained in interaction design examines a proposed design against a list of principles, guideline or heuristics.

- The term heuristics was introduced by Nielsen in the 90s as a generic expression to cover all sorts of principles and guidelines. He explained that he introduced the term "heuristics" because their similarity to "rules of thumb".

- There are many sets of guidelines and principles available (e.g. Benyon´s principles, or DIFI´s assessment criteria).

Heuristics Evaluation

Benyon's Principles:

A set of 12 principles (grouped into 3 broader categories) can be used as evaluating criteria:

Learnability
1. Visibility
2. Consistency
3. Familiarity
4. Affordance

Effectiveness
5. Navigation
6. Control
7. Feedback
8. Recovery
9. Constraints

Accommodation
10. Flexibility
11. Style
12. Conviviality

Heuristics Evaluation

DIFI Criteria:

- Difi aims for the improvement of digital public services by evaluating government websites.

- Evaluations of approximately one hundred randomly selected websites are carried out every other year.

- Difi´s assessment criteria are based on web standards, guidelines, best practice and political guidelines.

There are 32 criteria grouped into the following six categories:

- o Availability: how easy is it to find?
- o Credibility: how trustworthy is it?
- o Security: how safe is it to use?
- o Functionality: how well does it work?
- o Usability: how simple is it for anyone to use?
- o Support: how easy is it to get help?

Source: https://kvalitet.difi.no/kriteriesett

difi

Direktoratet for forvaltning og IKT

UiA University of Agder Norway

Cognitive Walkthrough

- A person trained in interaction design examines a proposed design by stepping through the cognitive tasks that must be carried out .
- A rigorous paper-based technique for evaluating design and interacting logic
- Three key prerequisites for this evaluation:
  - An understanding of the people who are expected to use the system and their goals.
  - A set of concrete scenarios.
  - A complete description of the system interface e.g. wireframes and navigation maps or a prototype.
- The evaluator aims to answer:
  - Will people know what to do at each step?
  - If people do the right thing, will they know that they did the right thing, and are making progress towards their goal?
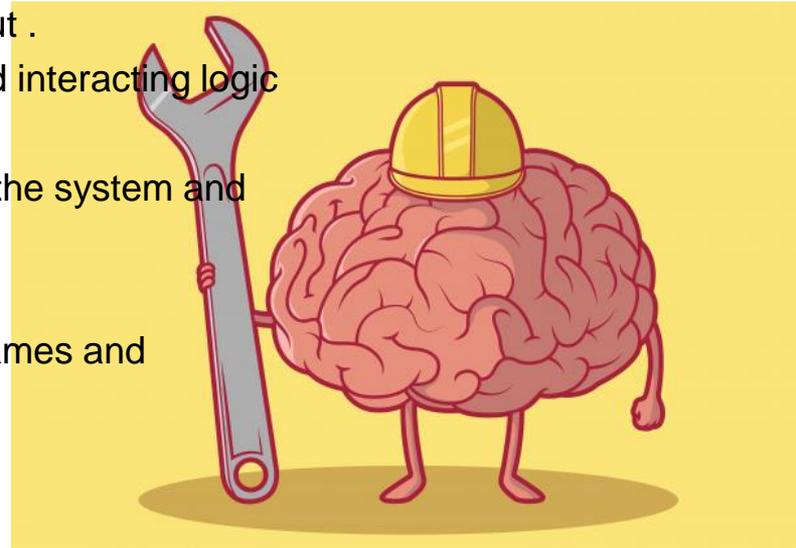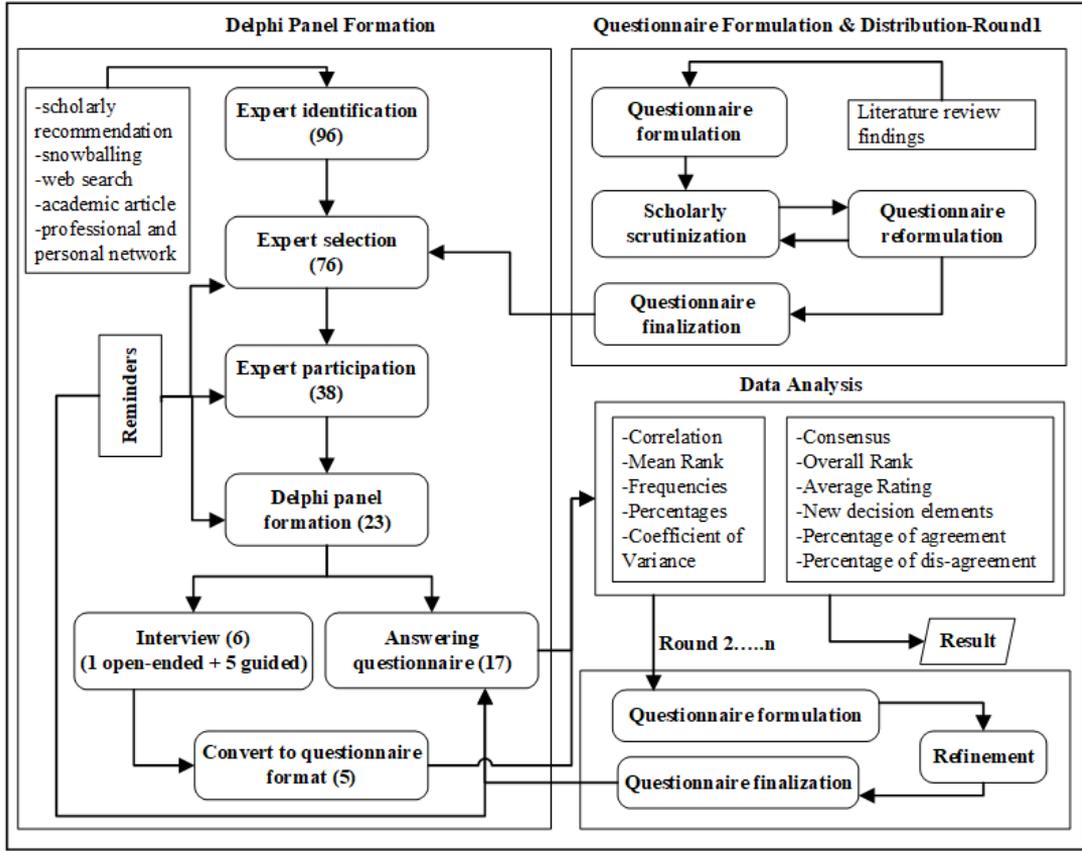
Expert Evaluation: Key Considerations

- Unless there is no alternative, **you should not evaluate your own designs**. It is extremely difficult to ignore your knowledge of how the system works, the meaning of icons or menu name.
- **Multiple evaluators are best:** Each evaluator is only uncovering some of the problems (often around 30% for more obvious issues) so having multiple evaluators will generate both more problems and identify overlapping problems.
- **Combining** task scenarios with heuristics is a widely used approach.
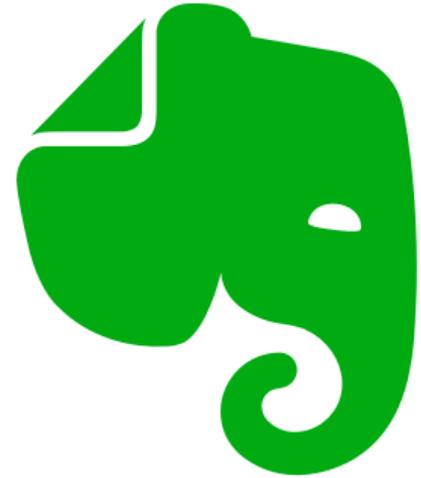
Find the simplicity of the demonstrated Delphi study process model:

- Take 10 minutes

- Is it a simple process model: why?

- Is it a complex process model: why?

Delphi Panel Formation

- scholarly recommendation
- snowballing
- web search
- academic article
- professional and personal network

Expert identification (96)

Expert selection (76)

Reminders

Expert participation (38)

Delphi panel formation (23)

Interview (6) (1 open-ended + 5 guided)

Answering questionnaire (17)

Convert to questionnaire format (5)

Questionnaire Formulation & Distribution-Round1

Questionnaire formulation

Literature review findings

Scholarly scrutinization

Questionnaire reformulation

Questionnaire finalization

Data Analysis

-Correlation
-Mean Rank
-Frequencies
-Percentages
-Coefficient of Variance

-Consensus
-Overall Rank
-Average Rating
-New decision elements
-Percentage of agreement
-Percentage of dis-agreement

Round 2.....n

Result

Questionnaire formulation

Refinement

Questionnaire finalization

UiA University of Agder Norway

Participant-based Evaluation

- No substitute for involving real people in the evaluation.
- There are many ways to involve people in evaluations:
  - sitting with participants as they work through a system
  - leaving people alone with technology and observing what they do through a two-way mirror
- It is needed to be ensured that
  - the people involved are representative of the prospective system **users**
  - the tasks that they are asked to achieve are related to specific **scenarios**
  - The evaluation is performed in **settings** that are as close as possible to the ones where use is expected to happen.
- Two main approaches:
  - Cooperative Evaluation
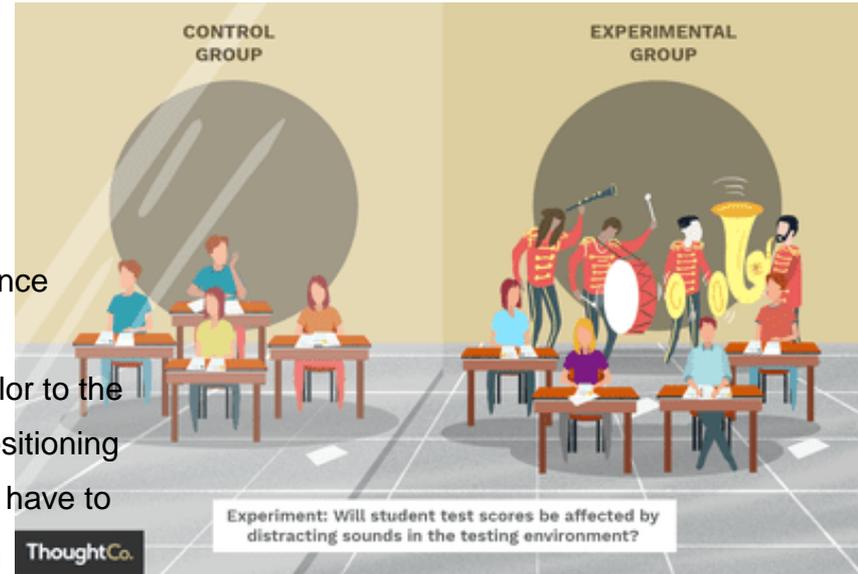 - Controlled Experiments

Cooperative Evaluation

- Participants are not passive subjects but work as co-evaluator

- Specific tasks given to participants based on the scenarios developed.

- Encouragement of participants to comment as they use the system.

- Interview with participants at the end.

- Performed when possible at the natural use setting ensuring "ecological validity".

- Audio-visual approach can be taken for explanation.

Steps for cooperative evaluation

| #	
1	Using the scenarios prepared earlier, write a draft list of tasks.
2	Try out the tasks and estimate how long they will take a participant to complete.
3	Prepare a task sheet for the participants.
4	Get ready for the test session.
5	Tell the participants that it is the system that is under test, not them; explain and introduce the tasks.
6	Participants start the tasks. Have them give you running commentary on what they are doing, why they are doing it and difficulties or uncertainties they encounter.
7	Encourage participants to keep talking.
8	When the participants have finished, interview them briefly about the usability of the prototype and the session itself. Thank them.
9	Write up your notes as soon as possible and incorporate into a usability report.

Controlled Experiments

- Designer is interested in particular design features:
  - For example, comparing designs to find the better one.
- An approach inspired by natural sciences.
- Key experiment design aspects are identifying:
  - independent variables: what it is that you are looking at
  - dependent variables: how you are going to measure the difference
  - confounding variables: things that might interfere experiencing.
- For instance, suppose you want to evaluate the impact of font color to the number of errors (dependent variable) or, the impact of button positioning to task completion time (dependent variable). In such cases, you have to ensure that your results will not be influenced by different lighting conditions, or learning effects, or different prior knowledge, etc.

CONTROL GROUP

EXPERIMENTAL GROUP

Experiment: Will student test scores be affected by distracting sounds in the testing environment?

ThoughtCo.

Controlled Experiments

- A controlled experiment will result in some quantitative data. This data can then be analyzed using statistics, for example, methods for hypothesis testing.
- Low fidelity prototypes cannot support timing studies and collection of quantitative data in general.
- A controlled experiment about user interfaces can be done in labs, or natural settings using simple equipment (i.e.,stopwatch) or specialized equipment for measuring specific variables. For example, eye tracking can capture gaze patterns (where a user looks at and for how long).
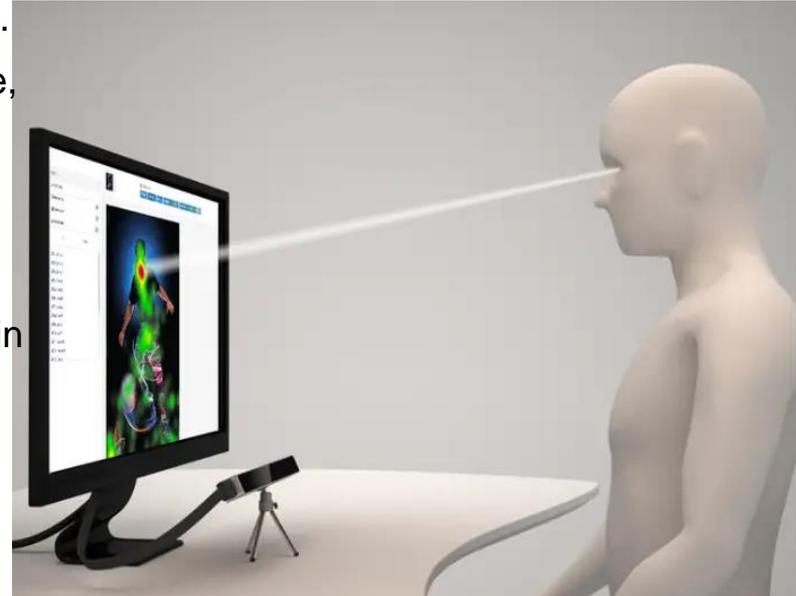
Participant Evaluation: Key Considerations

- Almost any degree of participant testing can reveal useful insights, but care must be taken in generalizing from a small **number of users**.
  - At least 3, better 5 participants – this number can be too low if you aim for statistical hypothesis testing
  - If you have a heterogeneous set of users, then you will need to run 3–5 people *from each group through your tests.*
- With the arrival of internet connectivity, people can participate in evaluations **without being physically present**.
- It is important to ensure that the participants are **representative** of the prospective users.
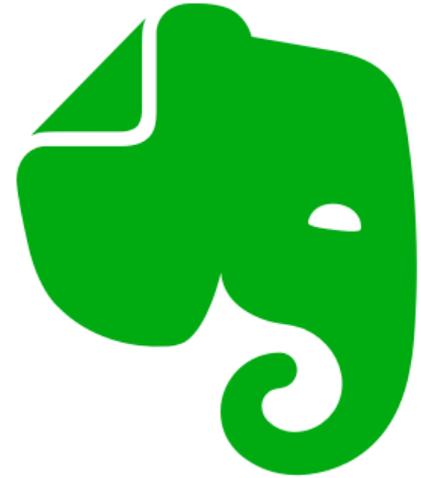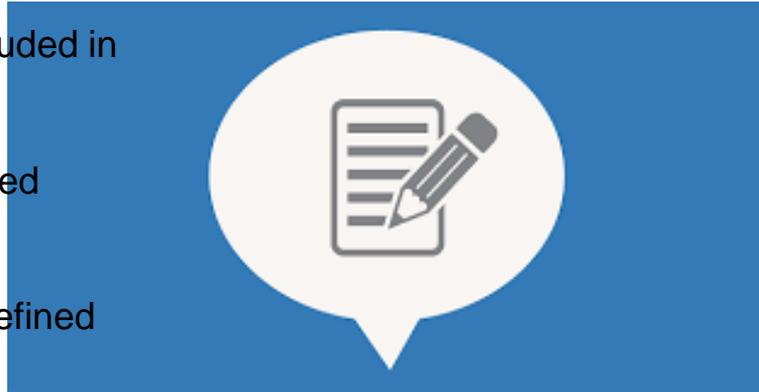
Data Analytics

- Can be gathered and analyzed when a system or service is implemented.
- It provides insights on
  - system performance and
  - the behaviors of individuals interacting with systems and services
- It provides designers with
  - data visualization and
  - tools to manipulate and analyze data
- Through data analytics individual and group activities can be examined
- By understanding the user behavior, Data analytics changes the nature of interaction software development – deploying new versions of software
- Performance can be compared and measured between interfaces by data analytics – used to refine UX of commercial websites.

Summary

- Expert-based methods are good for picking up UX issues quickly,

- But can miss detailed issues that are difficult for real users to find

- To get real feedback, participant-based methods must be included in the development process.

- Both types of evaluating method can be conducted in controlled and/or 'in the wild' environment.

- If real users are inaccessible, personas can be utilized with defined role types.

- When the system is implemented, data can be gathered and analyzed for further evaluation.

Thank you! Questions?

UNIVERSITY OF AGDER