NTNU

# Introduction to User-Centered Design

Lecture 9 – Evaluation part 1

Eleftherios Papachristos

Associate Professor at the Department of Design (NTNU)

# Today's lecture

## What will you learn?

## You will learn about evaluating your design

- Types of evaluation techniques
- Test planning
- Practical issues

NTNU

# Outline of this lecture

**NTNU**

**Part 1: Basic concepts of evaluation**

   – Definitions

   – Reasons to evaluate

   – Planning evaluations

**Part 2: Methods (What)**

   – Evaluation methods

   – Usability vs User testing

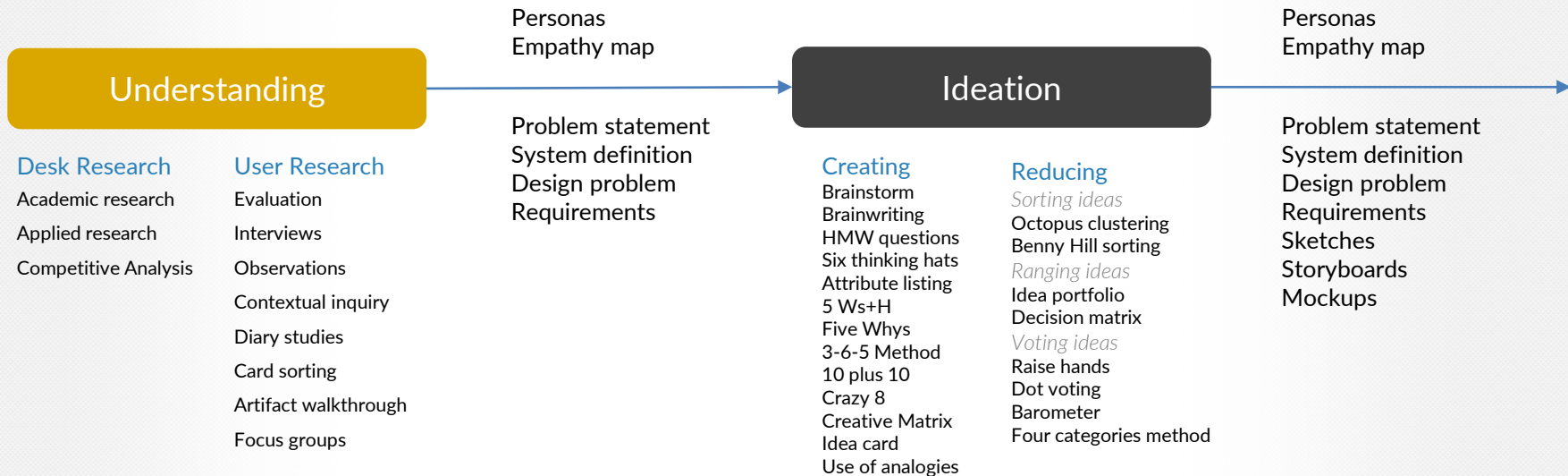   – Usability vs UX

**Part 3: Practical Considerations (How)**

   - Tasks

   - Moderator roles

   - Test environment

   - Equipment

**Part 4: Recruitment (Who)**

**Part 5: Results**

# Where are we in the process

Personas
Empathy map

Problem statement
System definition
Design problem
Requirements

## Understanding

## Ideation

Personas
Empathy map

Problem statement
System definition
Design problem
Requirements
Sketches
Storyboards
Mockups

### Desk Research
Academic research
Applied research
Competitive Analysis

### User Research
Evaluation
Interviews
Observations
Contextual inquiry
Diary studies
Card sorting
Artifact walkthrough
Focus groups

### Creating
Brainstorm
Brainwriting
HMW questions
Six thinking hats
Attribute listing
5 Ws+H
Five Whys
3-6-5 Method
10 plus 10
Crazy 8
Creative Matrix
Idea card
Use of analogies

### Reducing
*Sorting ideas*
Octopus clustering
Benny Hill sorting
*Ranging ideas*
Idea portfolio
Decision matrix
*Voting ideas*
Raise hands
Dot voting
Barometer
Four categories method

# Where are we in the process

Personas
Empathy map

```
                              ┌─────────────────┐              ┌─────────────────┐
                              │                 │  Prototype(s)│                 │
        ───────────────────►  │   Prototyping   │  ──────────► │   Evaluation    │
                              │                 │              │                 │
                              └─────────────────┘              └─────────────────┘
```

Problem statement
System definition
Design problem
Requirements
Sketches
Storyboards
Mockups

Type of prototype
Fidelity
Functionality
Physical prototype
Prototyping software

Type of evaluation
Planning
Recruitment
Study preparation
Process
Data analysis

Part 1

**1** Basic Concepts of evaluation

# What is Evaluation

Basic concepts of evaluation

*What do we mean by evaluation?*

It is the process of reviewing, trying out or testing a design idea, a piece of software, a product or a service to discover whether it meets some criteria.

Other names for this phase of the UCD process:

Testing, User testing, User research, User involvement, Usability testing

# What is Evaluation

Basic concepts of evaluation

*What does this mean practically?*

- It means watching people interact with your design by asking them to do **something** (complete a task).

- Your goal is to recruit appropriate users and select tasks that will guide them to aspects of your design you want feedback

- Afterwards you collect and analyse test data that will allow you to summarize the test and guide your further design process

# Why do we evaluate?

It is called User-Centered Design.

When did we involve users/humans? Did we?

## During the understanding phase

- Indirectly through desk research
    or
- Directly by user Observations/interviews

Is that enough?

# Why do we evaluate?

In order to understand :
- users
- their needs
- and their understanding of your product

**you should involve them as much as possible in all phases of the UCD development process**

In practice, this usually means during the understanding and evaluation phase
but

Don't forget that you should iterate multiple times
(involving users in those phases repeatedly)

# Why do we evaluate?

You can't assume how people will use your design (you must see them use it)
- It is often a humbling experience
- An aha moment
- Can be frustrating

Observing people using your design is the only way to learn where the issues are and how to fix them.

User feedback is priceless

It is cheaper to find issues with the design in this phase before going into development
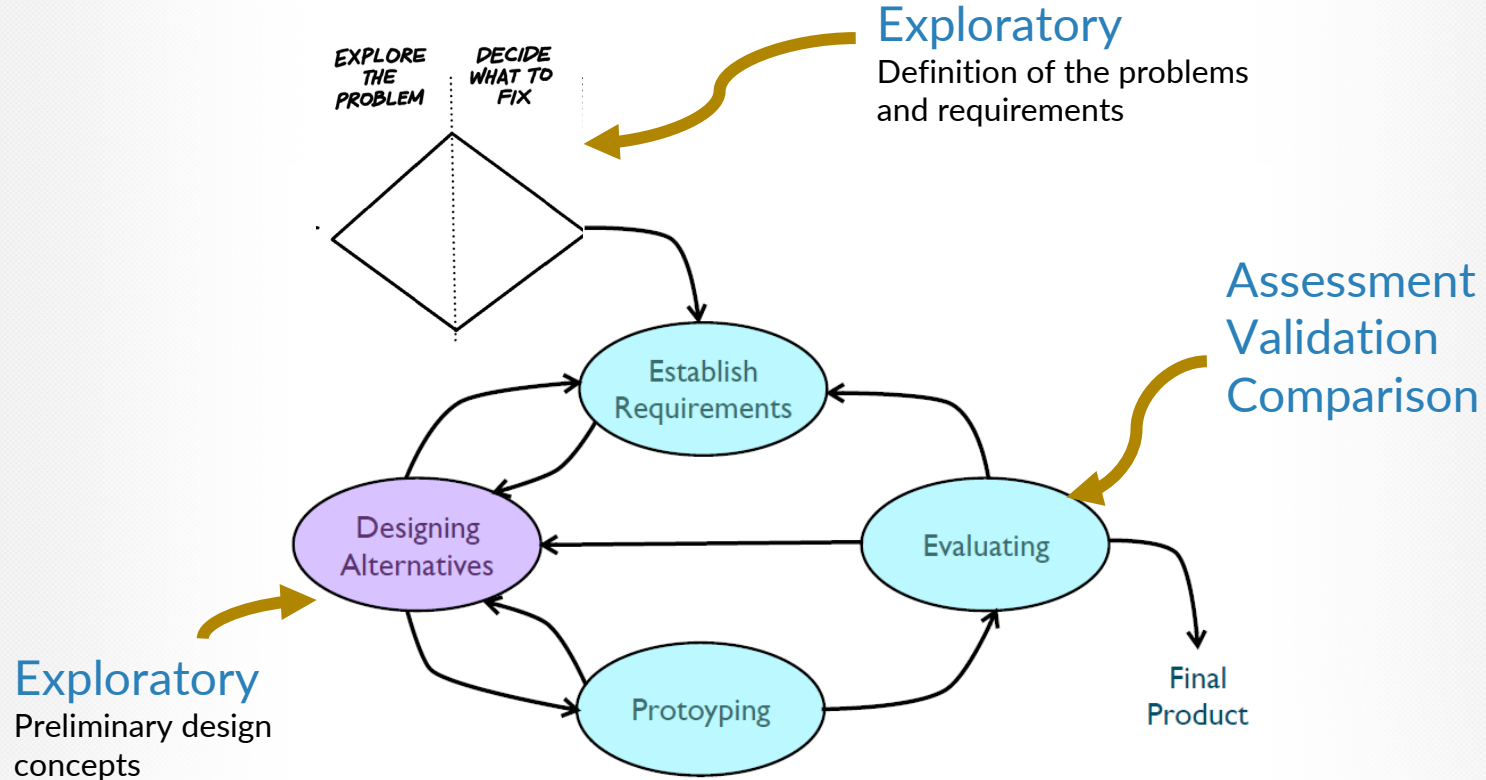
# Reasons to evaluate?

- Exploratory
  - Definition of the problems and requirements
  - Preliminary design concepts
- Assessment
  - Evaluating the usability of lower-level operations and aspects of the product
- Validation
  - Test against established benchmarks
  - Check if a problem has been solved
- Comparison
  - Compare two or more designs,
  - Compare your product with a competitor's.
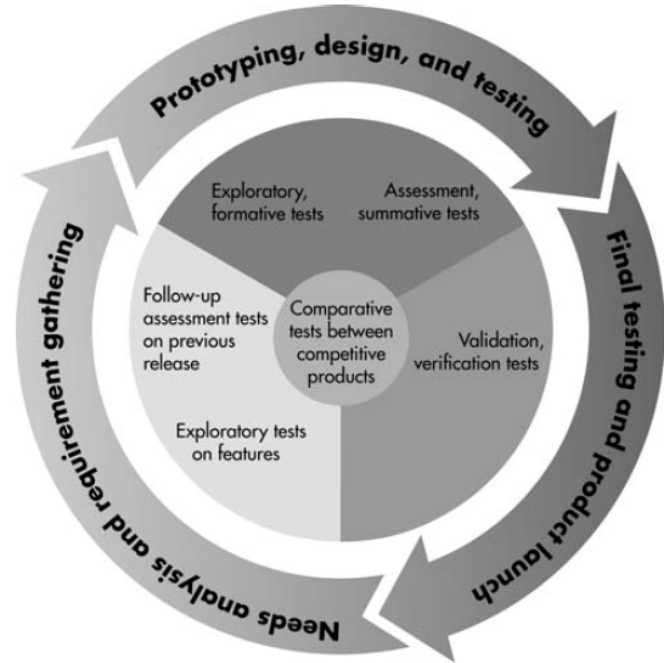
# Reasons to evaluate?

Basic concepts of evaluation

# Reasons to evaluate?

Basic concepts of evaluation

**Goal :** improve the quality of an interface by finding flows in it

**The main reason is to learn**, not to prove that our design is the best
- Failure is good (as long as we learn from it)

**Test our assumptions** about our design
- Not only what is problematic but also why

Check **if we have met the problem statement** ( or system definition)

# Evaluation Plan

Evaluation process

**(Why)** **Purpose, goals, and objectives of the test**
- Exploration, assessment, validation, comparison

**(What)** **Method** (test design)
- Formative or summative evaluation techniques

**(How)** **Tasks, Moderator role, Test environment, Equipment, and logistics**
- Practical testing consideration

**(Who)** **Participant recruitment**
- Who and how many should you involve in the evaluation

**(Results)** **Data to be collected and evaluation measures**

Part 2

**2** Methods (what)

# Evaluation methods

Table 14.1: Comparison of evaluation methodologies

| Method | Formative or summative | State of your product | Goal | Resources required |
| --- | --- | --- | --- | --- |
| Heuristic evaluation | Formative | Low to high fidelity | Identify violations of known usability guidelines | Low |
| Cognitive walkthrough | Formative | Low to high fidelity | Identify low-hanging issues early | Low |
| Usability testing in-person | It depends | Any stage | Identify usability issues | Medium |
| Eye tracking | Summative | Very high fidelity to launched | Identify where users look for features/information | High |
| RITE | Formative | Any stage | Iterate quickly on a design | High |
| Desirability testing | Summative | Very high fidelity to launched | Measure emotional response | Medium |
| Remote testing | Summative | Very high fidelity to launched | Identify usability issues across large sample | Low to High |
| Live experiments | Summative | Launched | Measure product changes with large sample of actual users | High |

Baxter, Courage, Caine **Understanding your Users**

# Formative  vs Summative methods

## Formative evaluations

Early in the product development life cycle to discover insights and shape the design direction. They typically involve user testing with low-fidelity mocks or prototypes.

## Summative evaluations

Typically done toward the end of the product development life cycle with high-fidelity prototypes or the actual final product to evaluate it against a set of metrics (e.g., time on task, success rate).

Baxter, Courage, Caine **Understanding your Users**

# Usability vs user testing

Many people use Usability testing and user evaluations interchangeably

**Evaluation IS NOT always Usability testing**

Usability testing is a type of an evaluation method

But

Not all evaluations are Usability tests

# Usability

Usability is a quality attribute that assesses how easy user interfaces are to use. The word "usability" also refers to methods for improving ease-of-use during the design process.

Usability is defined by 5 quality components:

**Learnability:** How easy is it for users to accomplish basic tasks the first time they encounter the design?

**Efficiency**: Once users have learned the design, how quickly can they perform tasks?

**Memorability**: When users return to the design after a period of not using it, how easily can they reestablish proficiency?

**Errors**: How many errors do users make, how severe are these errors, and how easily can they recover from the errors?

**Satisfaction**: How pleasant is it to use the design?

https://www.nngroup.com/articles/usability-101-introduction-to-usability/

# Usability

Definition

Usability is concerned with the "effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments" (ISO 9241-11).

Question to answer:  Can the user accomplish their goal

# Guerilla Usability test (example)

https://www.youtube.com/watch?v=0YL0xoSmyZI

# Usability

**Utility** = whether it provides the features you need.

**Usability** = how easy & pleasant these features are to use.

**Useful** = usability + utility.

Therefore

A design can be usable and at the same time not useful

https://www.nngroup.com/articles/usability-101-introduction-to-usability/

# Usability – Utility (example)
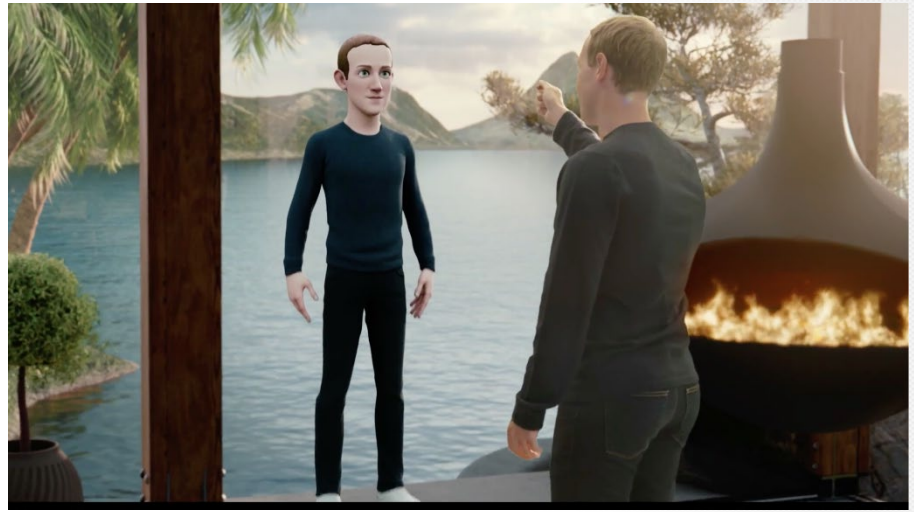
Facebook Metaverse

**Usability** = Is it working well (usable)?

**Utility** = Is it useful? Do people want/need this?

# Usability – Utility (example)

Google glass

**Usability** = Is it working well (usable)?

**Utility** = Is it useful? Do people want/need this?

# Usability vs User Experience

**Usability** = Usability is the measure of ease-of-use.

**User Experience (UX)** = User experience is a much broader.

*"UX includes all aspects of the end-user's interaction with the Company, its services and products"*

Don Gorman and Jakob Nielsen

It includes the users' thoughts, feelings and attitudes about your application website or interface. It involves the users' perception of usability

It is more subjective

https://qualitestgroup.com/insights/blog/ux-vs-usability-whats-the-difference/

# Usability vs User Experience

**Usability** = focused on users' goals.

**User Experience (UX)** =  focused on the entire experience.

Since usability has clear goals, requirements can be developed and then used in design and testing.

Since UX is subjective, what one user finds appealing may not be as appealing to another.

**Usability** = an the user accomplish their goal?

**User Experience (UX)** =  Did the user have as delightful an experience as possible?

# Example?

## Connectedness example

**Problem statement**(s)

How can we help grandparents increase their interaction with remote family members through fun activities on weekends?

**Usability** = important (Especially since our research showed elderly are not tech-savvy)

**But not enough** = what is missing?

- Is it fun?
- Will they use it?
- Will it bring them closer?

So, what do we do? Do we skip usability testing?

**Problem statement**(s)

How can we help grandparents increase their interaction with remote family members through fun activities on weekends?

**No. Usability is still important**

After usability testing we may:
- interview (debrief) participants and ask about utility
- Ask about the intention to use and ask WHY
- We may use a fun (game satisfaction) questionnaire

# User studies

**Goal** : to gain an understanding of people's behaviours, needs and motivations.

**How** :

- Through observation techniques

- You create scenarios and let users accomplish tasks (without your interference) while you observe them.

- Using think aloud protocol

# User studies

evaluation methods

- Very efficient to help you uncover unmet needs and assess your assumptions.

- **Very flexible method** (allows you to be broad, but, if necessary, you can go deep into an area of interest)

- You may also observe usability issues, but it is not the main goal

- **They do not require very large numbers of participants** (though more is better)

- Depend on experienced moderators, to ask open-ended questions and not lead the witness by showing agreement or displeasure with feedback provided by a study participant.

# Think aloud protocol

A think-aloud protocol is the process of having participants speak what they are thinking as they complete a task.

- Using a think-aloud protocol, you can get an understanding of why the user is doing what they do with your design

- Before asking participants to think aloud, it is helpful to provide an example, preferably one that reflects what they will be working with

- **This can feel unnatural, and they may forget to do it** (you must kindly ask them to continue doing it if they stop)

# Think aloud protocol

Example

**As you participate today, I would like you to do what we call "think out loud." What that means is that I want you to say out loud what you are thinking as you work. Let me show you what I mean. I am going to think out loud as I try to send a text message on this mobile phone. OK, I am looking at the home screen of the phone. I would expect there to be an icon that says "text" or "messages" right here on the home screen, but I don't see that here. I'm confused by that. I'm going to look in other places for it. I think I'll try to click on this icon of a callout bubble because that seems like it might be related to talking, like "I talk, the other person talks." Okay, now that that is open, I see an icon that says "text messages." I'm glad to see that there because that is what I was expecting to see. Do you see what I mean about thinking out loud? Now let's have you practice by telling me out loud how you would send an e-mail on this phone**

Baxter, Courage, Caine **Understanding your Users**

# Think aloud protocol

Ask participants **to talk aloud about what they are thinking** as they interact with a product **rather than their opinion about a product**.

- what they are looking at

- and what they expected

Also tell them you are not interested in how they think other people will feel about an interface.

E.g. "I don't think most people will think this color green is attractive. You should change it," redirect them to the task, and ask, "Tell me what you expected to see here."

Baxter, Courage, Caine **Understanding your Users**

# Think aloud protocol

## Can be used in:

- User studies

- Usability studies

- Card sorting / Eye-tracking ...

- Most types of methods involving users

Very useful in helping you understand why but can feel unnatural to participants. Also, it can change how they would normally use a design. You ask them to be more reflective and rationalise their behaviour

# Think aloud protocol

- **Retrospective Think Aloud** (RTA)
  The moderator asks participants to retrace their steps when the session is complete. Often participants watch a video replay of their actions, which may or may not contain eye-gaze patterns.

- **Concurrent Probing** (CP)
  requires that as participants work on tasks—when they say something interesting or do something unique, the researcher asks follow-up questions.

- **Retrospective Probing** (RP)
  Requires waiting until the session is complete and then asking questions about the participant's thoughts and actions. Researchers often use RP in conjunction with other methods—as the participant makes comments or actions, the researcher takes notes and follows up with additional questions at the end of the session.

https://www.usability.gov/how-to-and-tools/methods/running-usability-tests.html

# Usability test

**Usability tests are User studies with super focus on how easy it is to use our designs**

- It also involves observing real users use your design without prompting or help.

- Users are asked to complete tasks

- They are usually observed by a or researcher to see where they encounter problems and experience confusion.

Very similar to user testing. The difference is in how the tasks are formulated what is measured and what the focus is

# Usability testing methods

evaluation methods

## Many variations:

- **Heuristic Evaluation**

- **Cognitive walkthrough**

  *Inspection Methods*
  *Expert methods*

  Next lecture

- **RITE** (rapid iterative testing and evaluation)

- **Remote or in person tests**

- **Lab usability study**

- **A/B testing** (experiments)

# Usability testing methods

Usability testing is the systematic observation of **end users** attempting to complete a **set of tasks** with your product based on **representative scenarios**.

Usually participants interact with your product (e.g., paper prototype, low-or high-fidelity prototype, final product)

as they **think aloud** and user performance is evaluated against **metrics** such as:

- task success,

- time on task,

- Number of errors

-  conversion rate

**Several participants** are shown the same product and asked to complete the same tasks in order to identify as many usability issues as possible.

# Lab usability study

**User test happens in a dedicated testing space** (Lab).
You can create your own impromptu lab with a conference room, laptop, screen recording software, and a video camera.

**The Lab setting probably does not match the user's environment:**

- high-end equipment

- fast Internet connection,

- looks like an office,

- is devoid of any distractions (e.g., colleagues, spouses, or kids making noise

- and interrupting you).

**Lack ecological validity but offers control and  a consistent experience to all participants**

# In the Field usability study

To increase the ecological validity of your study you can conduct it in the field.

- If your product will be used at home, you could conduct the study in the participants' homes

- This will give you a better sense of how people will use your product in the "real world."

- Less control

# Remote usability study

It is not always possible, to conduct evaluations with participants in person.
(e.g. participants with specific demographics)

## Remote studies

- Allows you to collect feedback from a much larger sample size in a shorter period of time, and no lab facilities are needed.

- Not possible for physical design evaluations

You can Use online services to conduct evaluations with their panels (e.g., *UserZoom, UserTesting.com, Amazon Turk, Prolific Academic*).

You can use a service video conferencing tools to remotely connect to the participant in his or her environment (e.g., home, office) while you remain in your lab or office.

# Rapid Iterative Testing and Evaluation

Formative evaluation method developed by Microsoft's game division to quickly address issues that prevented participants from proceeding in a game and evaluating the remaining functionality

- RITE is designed to *quickly* identify any large usability issue that is preventing users from completing a task or does not allow the product to meet its stated goals.

- Should be conducted early in the development cycle with a prototype

# User vs Usability Test

## Usability test

- Formal regarding task
- Emphasis on Effectiveness and Efficiency (by measuring tasks completion rates. time. error rates)
- Data analysis emphasis on usability metrics and severity of errors

## User test

- Less formal regarding task (include a free play task)
- Apart from usability it tries to assess other factors too more emphasis on other qualities (e.g. utility enjoyment, fun, intention to use, if it meets problem statement)
- Analysis emphasises on both usability metrics and assessment of other qualities (additional questionnaire, debriefing interview)

Both can be in the lab, in the field, or remotely

Both can use think aloud protocol

Both can have a debriefing interviews
(were the emphasis is where the difference is)

Part 3

**3** **Practical considerations** (how)

- Tasks
- Moderator roles
- Test environment
- Equipment

# Formal testing

Participants

Environment

Procedure
(Tasks)

Digital artifacts

Measurement tools

Researcher

# Sources of systematic errors

overview

In order to eliminate systematic errors, we have to identify where they originate

- Participants

- Researchers

- Digital artefacts

- Environment

- Procedure

# Participants (Study Subjects)

Participants carry their own experiences and expectations, and have different characteristics
- **They may not be representative of the population** (wrong sampling)
- **They get tired if we keep them a long time** (fatigue effect)
- **They may respond in unexpected ways** (e.g. desirability bias)

**Tips.**

**Recruit participants that fit your target group** (representative of the population)

**Try to minimize the fatigue or stress either by keeping experiments short or having breaks**
(Sometimes it makes sense to introduce irrelevant activities during breaks)

# Researchers

Make attempts to not influence participants

**However, we can also unintentionally influence results by just being there** (the way we phrase questions, body language, facial expressions)

Tips.

- The same researcher always communicates with the participants
- **The same instructions are be provided every time** (a script or a video is a good solution)
- Always be polite but never comment on participants behavior
- **Always have a talk after the experiment** (debriefing session)
- Have a script

# Digital artifacts (mediums)

The digital artifacts our participants will interact with might have characteristics that can influence the experiment (screen sizes, keyboards, brands, etc.)

**Tips.**

- Carefully select the digital artifacts the participants will interact with and try to remove any influential characteristics
- Keep the same digital artifacts throughout the experiment
- Make sure the perform the same way all the time

# Procedure

- The framing of the tasks the participants will receive can influence the results of an experiment.
- The practice gained in one task will influence the rest (practice effect)
- The experience they will have from a task will influence the rest (carryover effect)

Tips.

- Provide clear and easy to understand tasks and scenarios to the participants
- Always randomize or counterbalance the sequence of the tasks

# Tasks and scenarios

**(Tasks)** In order to observe participants, you need to give them something to do.

Examples
- Buy a pair of plain ticket to Athens in December for less than 1500 kroners

- Use the website to find a movie you'd be interested in seeing on Sunday afternoon

**(Scenarios)** Instead of simply asking users to do tasks with no explanation, it's better to situate the request within a short scenario that sets the stage and provides a bit of explanation and context.

Example
- "You're planning a vacation to Athens Greece, December 18 – December 27. You need to buy both airfare and hotel. Go to the website (url) and see who has the best deals."

https://www.nngroup.com/articles/task-scenarios-usability-testing/

# Tasks

procedure

## Before writing task and scenarios

Come up with a list of general user goals that visitors to your prototype may have.

Ask yourself: What are the most important things that every user must be able to accomplish on the site?

Examples of goals

- Find articles on a specific topic
- Sign up for a seminars
- Learn about services

Apart from user goals also think about what features/sites/ functionality of your prototype you would like feedback on

# From goals to tasks

## Make the Task Realistic

- User goal: Browse product offerings and purchase an item.
- Task: Purchase a pair of orange Nike running shoes. **(poor task)**
- Task: Buy a pair of shoes for less than $40. **(better task)**

Asking a participant to do something that he wouldn't normally do will make him try to complete the task without really engaging with the interface.

## Make the Task Actionable

- User goal: Find movie and show times
- Task: You want to see a movie Sunday afternoon. Go to (url link) and tell me where you'd click next. **(poor task)**
- Task: Use (url link) to find a movie you'd be interested in seeing on Sunday afternoon. **(better task)**

It's best to ask the users to do the action, rather than asking them how they would do it.

https://www.nngroup.com/articles/task-scenarios-usability-testing/

# From goals to tasks

## Avoid Giving Clues and Describing the Steps

- User goal: **Look up grades.**
- Task: **You want to see the results of your midterm exams. Go to the website, sign in, and tell me where you would click to get your transcript. (poor task)**
- Task: **Look up the results of your midterm exams. (better task)**

Step descriptions often contain hidden clues as to how to use the interface. For example, if you tell someone to click on Benefits in the main menu, you won't learn if that menu label is meaningful to her. These tasks bias users' behavior and give you less useful results.

## Include a free play period without constrains or tasks

- Say to users: **play around with the prototype for a while, doing anything you would like to, and feel free to talk aloud while you are doing this**

https://www.nngroup.com/articles/task-scenarios-usability-testing/

# The test team

user  facilitator  computer  scriber & more

# Before the test

## Clarify Goals
Discuss what the purpose of the test is and agree on what aspects of the design you will focus testing

## Define Tasks
Discuss and identify and write down the user tasks and scenarios you will give to the users. Write test script.

## Additional measures / interview
Find questionnaires or write interview guid for debriefing sessions

## Run a pilot
It is always important to test the test among you before inviting users

# Start of the test

## Briefing

Explain who you are you and why you are doing this test
In general what will they have to do and how long will it take
**Think-Aloud** ask them to say what they are thinking as they go.

## Task

Give the participants a **task** " log in your personal account and do a bank transfer" "log in and upload your assignment" "check the delivery date for your next assignment"

# During the test

- The user goes trough the task interacting with the interface.

- The user says out loud what is going on

- If the user ask questions he should ask to wait to the end (if you measure task completion time if not respond immediately).

- You simulate the feedback of the interface if not part of the prototype. (wizard of Oz or paper prototype)

- If the user gets to a missing page do not tell what would be there, ask what they expect to find

# After the test

## Debrief

- Ask for comments, thoughts or questions.
- The note taker should ask if he needs to fill in some gaps.
- **Reward** (coffee? Cookies? )

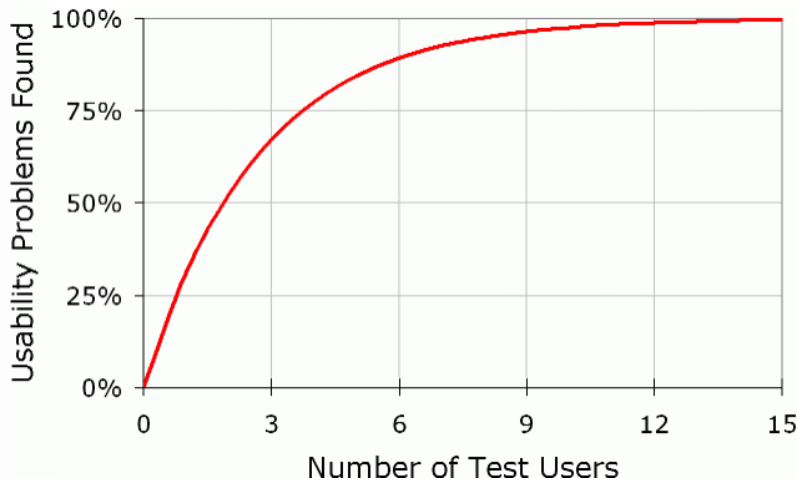- Consent form

Part 4

**4** **Recruitment** (who)

# Number of participants

**For usability tests (important this is not true for other test)**

- You don't do more than 5 test for each iteration.
- You will find most of the issues in the first few tests**.**



**Quantitative studies** (aiming at statistics, not insights): **you need 20-30 users** to get statistically significant numbers.

# Who should you recruit

**The closer to your end user the better**

You try to recruit through:

- Social media,

- Recruitment services

- Snowball recruitment

- Posters, adds, give incentives

**If recruitment is difficult**

- Use your personal network (convenience sample) but be aware and reflective about the shortcomings

- Especially for usability it is better to test with some people you know rather than no testing at all

# Who should you recruit

Sources of systematic errors

**Use your personas to identify target users**

If you cannot find appropriate target users: Use a convenience sample

- This is less problematic if you do a usability test
  (except if your design needs domain knowledge)

- More of a problem in a UX test

Still some test is better than no test but be careful about the results

Part 5

**5** Results (data collection)

# Usability metrics

**Usability metrics**

.

- **Time on task**: Length of time to complete a task.

- **Number of errors**: Errors made completing a task and/or across the study.

- **Task Completion rate**: Number of participants that completed the task successfully.

- **Satisfaction:** Overall, how satisfied participants are on a given task and/or with the product as a whole at the end of the study

- **Page views or clicks**: As a measure of efficiency, you can compare the number of page views or clicks by a participant against the most efficient/ideal path.

- **Conversion:** measure of whether or not participants (users) "converted" or successfully completed their desired task

# Debriefing

Sources of systematic errors

**Interviews and questionnaires**

.

**You can** (and you should) **have a short debriefing interview after a user or usability test**
Ask questions that would allow you to understand:

- Motivation
- Intention of future use
- Overall user experience

You can use post-study questionnaires if you need to assess some quality

– **Aesthetics Scale** (classic and expressive aesthetics)
– **Attractiveness Questionnaire** (attractiveness)
– **Attrakdiff2** (hedonic and pragmatic quality)
– **SUS Questionnaire** (usability)
– **Pleasure with Products Questionnaire** (pleasure)

# Result reporting

For usability tests use descriptive statistics for the metrics

- **Time on task**
- **Number of errors**
- **Completion rate**
- **Satisfaction**
- **Page views or clicks**
- **Conversion**

Such as simple Averages, Standard Deviations, and Bar charts to create an overview

# Result reporting

Sources of systematic errors

Use screenshots and describe the problems you have identified

Categorize and summarize problems based on severity

| | UES-HE Method | | URS-HE Method | | HE Method | | UT Method | |
|---|---|---|---|---|---|---|---|---|
| 1: Cosmetic | 15 | 30.0% | 11 | 30.6% | 8 | 33.3% | 10 | 23.3% |
| 2: Minor | 13 | 26.0% | 9 | 25.0% | 8 | 33.3% | 11 | 25.6% |
| 3: Major | 17 | 34.0% | 12 | 33.3% | 5 | 20.8% | 16 | 37.2% |
| 4: Catastrophic | 5 | 10.0% | 4 | 11.1% | 3 | 12.5% | 6 | 14.0% |
| Total real problem | 50 | 100.0% | 36 | 100.0% | 24 | 100.0% | 43 | 100.0% |



Figure 7: Track Listing. Test participants found the listing of tracks on individual albums informative. They also appreciated that they could hear samples of some of the tracks.

All test participants found it helpful that they could listen to 30-second samples of the tracks on the individual albums. See Figure 7. "This is a help for me. There are very few albums I would buy without listening to them first."

Four of the six participants commented that it was annoying that they could not hear samples of music from all albums. "When they finally have something I am interested in then why can't I hear it?" One participant noted that the newest albums seemed to be the ones you could not hear.

Test facilitator's comment: None of the test participants looked for an explanation. We have not been able to find one.

https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/

# Result reporting

## Categorize and summarize problems based on severity

| Cosmetic | This is the lowest priority and may only get fixed if there is time available. *To be categorized as a cosmetic problem, a problem has to only affect the website's appearance, and not the functions.* |
|---|---|
| Minor | As it is a minor usability problem, it should get fixed eventually, but it has a low priority. *To be categorized as a minor problem, a problem has to affect a function, but do not prevent the user to complete the intended task without any major delay.* |
| Major | A major usability problem should be prioritized over others. *To be categorized as a major problem, a problem has to affect a function and cause either major delay or prevention to completing the users intended task.* |
| Catastrophic | A system should not get released without fixing this problem. Can cause catastrophic consequences. *To be categorized as a catastrophic problem, a problem has to prevent all chances of completing the intended task. Such a problem shall have the biggest priority and need to be fixed immediately.* |

# Result reporting

## Categorize and summarize problems based on severity

| Severity | Problems | Description | frequency |
|---|---|---|---|
| Minor | P1 | The test user tries to click on a non-clickable bar on the frontpage (se rejser og saldo). | 5 |
| Minor | P2 | The test user oversees the 'Beregn rejse' function. | 3 |
| Minor | P3 | The test user wants to log in to do the task ("beregn rejse"). | 1 |
| Minor | P4 | The test user tries to find the function in 'spørgsmål og svar'. | 2 |
| Cosmetic | P5 | The test user can not find the 'Beregn pris'-button. | 2 |
| Cosmetic/ test-leader error | P6 | The test user does not type in the correct time and date. | 3 |

| Problem | Category | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1: No clear indication on where to mark allergies or any other change | Critical | X | X | X | X | X |
| 2: Old address, giving wrong delivery days | Serious | O | X | X | X | X |
| 3: Lack of a sorting system, when choosing meal box | Serious | X | X | X | X | X |
| 4: Hidden information about how to cancel subscription | Critical | X | O | X | X | O |
| 5: Confusing function, when clicking on 2 different buttons, that should have the same function when used. | Critical | X | X | X | X | X |

# Result reporting

Sources of systematic errors

## Time on task

| Task number | Description of task | Time in sec. |
|---|---|---|
| Task 1 | Find a womens bike price <2000 DKK. | 116 |
| Task 2 | Create an account. | 264 |
| Task 3 | Add an advert to your favourites. | 104 |
| Task 4 | Find a corner couch priced >5000 DKK. | 265 |
| Task 5 | Create an advert. | 226 |
| Task 6 | Find the rules for selling animals on DBA. | 153 |
| Task 7 | Find a phone number to get in touch with DBA. | 136 |

| Subject nr. | Task number and time (min.sec) | | | | | | Total time |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | |
| 1 | 3.34 | 2.58 | 3.51 | 2.43 | 0.38 | 4.27 | 19.15 |
| 2 | 2.16 | 1.15 | 4.11 | 3.15 | 0.42 | 2.55 | 17.40 |
| 3 | 3.06 | 0.54 | 0.48 | 1.55 | - | 1.10 | 9.22 |
| 4 | - | 3.03 | 1.20 | 7.32 | 0.50 | 1.49 | 16.04 |
| 5 | - | 1.55 | 2.19 | 2.01 | 0.18 | 0.20 | 7.50 |
| | | | | | | | |
| Avg. time | 2.59 | 1.57 | 2.32 | - | 0.37 | 1.56 | 14.02 |

# Reading material for this Lecture

## Reading
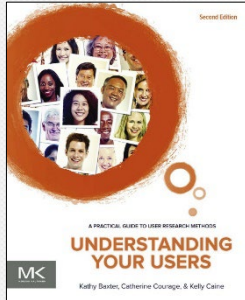
Baxter, Courage, Caine **Understanding your Users**

- **Chapter 14** Evaluation Methods

# For the Lab session

NTNU

Meet up in the Lab rooms to develop and develop a script for your prototype evaluation

- Decide on evaluation method

- Define scenarios and task

- Recruit people (if necessary invite each other)

- Run the test an right up the results

# Thank you

**Lefteris Papachristos**

Associate Professor, NTNU