# On measuring and measurements

# What does it mean to measure?

# What does it mean to measure?

To measure: “To assign a value to a variable”

# What does it mean to measure?

To measure: “To assign a value to a variable”

“Variable”

“Value”

“Assign”

# What does it mean to measure?

To measure: “To assign a value to a variable”

“Variable”

- (1) Effectiveness; (2) efficiency; (3) satisfaction

“Value”

“Assign”

# What does it mean to measure?

To measure: “To assign a value to a variable”

“Variable”

- (1) Effectiveness; (2) efficiency; (3) satisfaction

“Value”

- (1) per cent of population; (2) seconds; (3) rating on a scale

“Assign”

# What does it mean to measure?

To measure: “To assign a value to a variable”

“Variable”

- (1) Effectiveness; (2) efficiency; (3) satisfaction

“Value”

- (1) per cent of population; (2) seconds; (3) rating on a scale

“Assign”

Quality of the assignment

# Quality of a measurement

- Reliability
- Validity

# Quality of a measurement

- Reliability
- Validity

- What does this mean in our case?

# Common types of evaluation

- Heuristic based
  - Heuristic evaluation
- Task based
  - Formative usability testing
    - Participatory evaluation
  - Summative usability testing
  - Validation study
  - Comparison study

# Heuristic evaluation

1. Visibility of System Status
2. Match between system and the real world
3. User control and freedom
4. Consistency and Standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility and efficiency of use
8. Aesthetic and minimalist design
9. Help users recognize, diagnose and recover from errors
10. Help and documentation

*https://www.nngroup.com/articles/ten-usability-heuristics/*

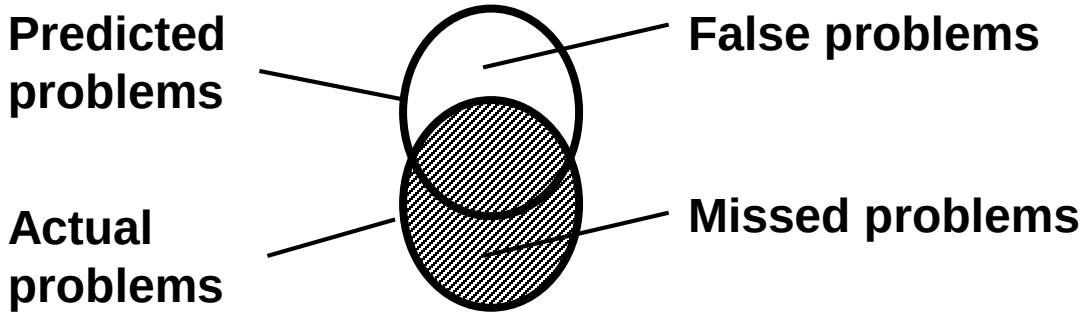
# HE: Procedure

- Each person looks at the interface alone
  - Screen by screen (or principle by principle)
  - Notes:
    - Found problem
    - What heuristic it relates to
    - Its severity
- Then the group assembles
  - One session leader, one note-taker
  - The found problems are discussed
    - Collapse problems
    - Rate severity

# Heuristic evaluation

- Can only find potential problems
- Why?
  - Since knowledge of the user population, their tasks and their context of use isn't (explicitly) used, or quite often, not even known.

**Heuristic evaluation (maybe! acc. to Laursen) has only 50% hitrate**



# CI: “First, identify core user groups”

- No agreed upon optimal way
- Best practice:
  - A simplified context of use seminar:
    - Gather key stake holders and some interested intended end users
    - Go through the CoU form and process
  - Result:
    - User groups in the form of personas and related info

# Formative testing

- Purpose:
  - To find potential usability problems
  - To establish a measure on effectiveness

# Summative testing

- Purpose:
  - To establish a measure on efficiency

# Parts of a test plan

- Purpose
- Research questions
- Participant characteristics
- Method (test design)
- Task list
- Test environment, equipment and logistics
- Test moderator role
- Data to be collected and evaluation measures
- Report contents

- Steps in a usability evaluation
  - *Define the goal of the test*
    - Create hypotheses ('Diagnostic' or 'Formative')
      - Define measurements
        - *"Productive time"*
        - *Occurrence of problems (define!)*
    - Evaluating ('Summative')
      - Against what?
        - *Set goals*
        - *Competing products*
        - *Expert performance*
        - *Existing system*

- *Define populations of*
  - Users
  - Environments
  - Tasks
- *Select samples*
  - Users
    - Participant selection
      - *Stratification + random sampling*
      - *If too costly*
        - Representativeness
    - Assignment (if more than one version of system)
      - *Random*
      - *Matched*
  - Tasks
    - Random sampling too costly
    - Representativeness
    - More than one task?
      - *Random assignment*
      - *Counterbalancing*
  - Environments
    - Differences between the real world and your measurement situation
    - The actual situation
      - *Experimenter effect*
      - *Expectancies*
    - What can be done?
      - *Blind procedures*
      - *Automation*

- *Choose things to register*
  - Governed by earlier choices
  - Two types of main things
    - Events
    - States
  - Remember!
    - Always operational definitions of everything

- ***At test time***
  - **Prepare**
    - **Cue cards (tasks)**
    - **Recording device**
    - **Test leader material**
  - **Welcome the person**
  - **Initial session one on one**
    - **Ethical information and consent forms**
    - **Ensure the participant that it is the system being measured**
    - **Pre-test questionnaire**
  - **Warm up tasks**
  - **Administer tasks according to plan**
  - **Do not interfere with test person unnecessarily and never in such a way that the validity of the measurements are endangered**
    - **Ethical rules must be adhered to, of course**
  - **Debrief**
  - **Back up copy of recording**

***Analysis***

Don’t forget to validate observers!

***Prepare report***