

# Week 12-3

# **Usability Testing**

SFWRENG 4HC3/6HC3 Human Computer Interfaces

*\* Slides adapted from previous and current instructors of COMPSCI/SFWRENG 4HC3/6HC3*

# Heuristic Evaluation: Practice

**Take 5-7 minutes** to conduct a heuristic evaluation for the McMaster CAS website (<https://www.eng.mcmaster.ca/cas/>), on the task for **finding course requirements and information** for a software engineering student.

- Familiarize yourself with the 10 heuristics before starting the task
- Walk through the website to complete the task, and take notes on the usability issues you notice

Submit your top 2 findings on Avenue (relating to specific heuristic)

# Week 12 Overview

- **Tuesday**

- ~~Introduction to Evaluation~~
- ~~Evaluation: Performance Modeling~~

- **Thursday**

- ~~Evaluation: Inspection~~
- ~~Evaluation: Heuristics~~

- **Friday**

- **Evaluation with Users**

# Evaluation with Users

**Different goals** (when, where, what), lead to **different methods**

- Controlled (laboratory) experiments
- Field experiments
- Field studies
- Qualitative usability studies/tests
  - Think aloud method

# Evaluation with Users: Overview

Kind of learning	Formative	Summative
Goals	Exploration	Evaluation
Type of data	More Qualitative	Quantitative
Level of control	Less	More
Formality	Generally Less	Generally more
Phase	Design/Prototype	Testing
Cost	Often cheaper	Often \$\$\$
User tasks	Relatively open	Assigned

# Evaluation with Users: Summative

- Concrete, **quantitative** measures of usability
  - Time to learn a feature
  - Use time for specific tasks
  - Features used (or not)
  - Error rates
  - Measures of user satisfaction
  - Comparison to prior/alternative versions, competitors
- **Results**

# Evaluation with Users: Formative

- **Qualitative** experiences of usability
  - What will they use this thing for anyway?
  - Trouble spots in completing tasks
  - Features found / not found
  - Reactions to design elements/decisions
  - Learning users' mental models
  - *Why* can't users do it?
- **Guidance**

# Evaluation with Users: Ethics Principles

- **Voluntary, Informed** Consent
  - Informed –enough to make consent meaningful
  - Voluntary –including right to stop at any time
- **Do No Harm**
  - Make it clear this isn't about them, it is about the product!
    - That is also the line between “regulated research” and a usability test/user study
  - Be careful and thoughtful about recordings
  - If things go wrong, consider how to end the test

# Evaluation with Users: Ethics Practices

- **Formal “Bill of Rights” and Consent Form**
  - Add specifics about the test content itself
  - Add measurements
  - Be sure to give an opportunity for questions
- Explain **Anything You Can**
  - Use debrief for things that would interfere with usability test/user study
  - Give test participants a chance to share comments, ask questions, etc.
- Culture of **Treating Participants Well**

# Evaluation with Users

**Different goals** (when, where, what), lead to **different methods**

- Controlled (laboratory) experiments
- Field experiments
- Field studies
- **Qualitative usability studies/tests**
  - Think aloud method

# Usability Test: Overview

A usability test is a “formal” method for **evaluating** whether a design is **learnable, efficient, memorable, can reduce errors, meets users’ expectations, etc.**

Users are not being evaluated; **the design is being evaluated**

# Usability Test: Participants

- Bring in real users
- Have them complete tasks with your design, while you watch (ideally with your entire team)
- Measure and record things
  - Task completion, task time, error rates
  - Satisfaction, problem points, etc.
  - Might use a think-aloud protocol, so you can “hear what they are thinking”

# Usability Test: Data

Use the data to

- Identify **problems** (major ones & minor ones)
- Provide **design suggestions** to design/engineering team
- **Iterate** on the design, repeat

# Usability Test: Planning

- Usually takes place in a usability lab or other **controlled** space
- Major emphasis is on
  - Selecting representative users
  - Developing representative tasks
- 5-10 users typically selected
- Tasks usually last no more than 30 minutes
- The test conditions should be the same for every participant
- Informed consent form explains ethical issues

# Usability Test: Environments



# Usability Test: Environments

Best practices for choosing the environment depends on **pragmatic considerations**, as well as **what you're looking for**:

- Do you want your **whole team** to be able to view?
- Do you want to be able to **review** a test?
- How important are **interruptions**?
- What are your **resources**?

# Usability Test: Piloting

- **Especially important** for usability testing
- Make sure your plan is viable
- All the corners are checked (your script, questionnaires, tasks, etc., all work)
- It is worth doing several to iron out problems before doing the main study
- **Ask colleagues/peers** if you can't spare real users

# Usability Test: Tasks

- A task is designed to **probe a problem**
- Tasks should be **straightforward** and require the user to **find certain items, or do certain operations**
- They can be more complex such as solving particular problems
- Sample tasks for a weather network web site:
  - What is the forecasted weather for Vancouver?
  - What is air quality in Los Angeles today?
  - What is the level of humidity in Hamilton?
  - What is the forecast for Ottawa for the upcoming weekend?

# Usability Test: Tasks

You are developing a user test for a new CAS web page. **What are some tasks that you might design for the user test?**

Task 1: Identify the instructor for 4HC3

Task 2: Find the e-mail address of the 4HC3 prof

Task 3: Find the admission requirements for the M.Sc. Program

Task 4: Find out the first day of classes next term

Task 5: Locate the requirements for being a Co-op student

# Usability Test: Tasks

Tasks can also be more complex, such as **solving a particular problem**

Complex tasks may tell you more, but users may get lost in particular details

# Usability Test: Number of Users

- The number is largely a **practical issue**
- Depends on:
  - Schedule for testing
  - Availability of participants
  - Cost of running tests
  - Will you try to publish the results as scientifically sound?
- Typical 5-10 participants
- Some experts argue that testing should continue until no new insights are gained

# Usability Test: Data Collection

Draw from requirements elicitation/user research methods

- Questionnaire
- Observation
- Think-Aloud
- Interviews

# Usability Test: Questionnaire

- Earlier in the term we discussed questionnaire design for gathering requirements
- In usability tests, questionnaires **typically focus on user experience goals** (e.g., satisfaction) and **consist primarily of closed questions**
  - Participants encouraged to leave their comments in space provided on the page, or in the margins
- More about designing closed questions well

# Usability Test: Questionnaire

**Likert-like scales** are used for measuring opinions, attitudes, beliefs.

**For example,** evaluating color on a web site can have the forms:

The use of color is excellent:  
(where 1 represents strongly disagree and 5 represents strongly agree)

1

☐

2

☐

3

☐

4

☐

5

☐

Strongly  
disagree

☐

Disagree

☐

Neutral

☐

Agree

☐

Strongly  
agree

The use of color is excellent:

24

# Usability Test: Questionnaire

**Steps** for designing Likert-like scales

1. Gather a pool of short statements about the features of the product that are to be evaluated
2. Create logical/conceptual groups
3. Decide on the scale (5-point/3-point/7-point)
4. Select items for the final questionnaire and reword as necessary

# Usability Test: Observation

- The majority of evaluations with users involve **some form of observation**
- Simple form of observation
  - User is given a set of tasks, and the evaluator simply watches the user

What do you watch? What do you do?  
What do you record?

# Usability Test: Observation

- **Unobtrusive observation:** be quiet, watch, understand
- **Don't explain, help or defend the design,** don't apologize, don't help out (hard to watch them struggle!)
- **Answer questions with questions:** can be an opportunity to enrich observation
  - User: "Do I have to click here?"
  - Observer: "What do you think will happen if you click there?"
- Only help to **overcome the limitations** of the prototype
  - Explain briefly and neutrally what would happen in the future system
  - Help them get un-stuck, and let them know it's a limitation

# Usability Test: Observation

- Sometimes **direct** observation can be obtrusive or impossible
- Alternative approaches
  - **Interaction logging**
    - Recording key presses, mouse buttons, interface changes
    - Difficulty: need to correlate specific action with the appropriate tasks and meaning (hard)
  - **Diaries/experience sampling**
    - What users did, when they did it, and what they thought about their interactions
    - Provide templates for users to fill in

# Usability Test: Think-Aloud

Gives insight into what the user is thinking

## **However:**

- Awkward/uncomfortable for subject
- May alter the way people perform their task
- Hard to talk when they are concentrating
- User's personality may not align with thinking aloud

# Usability Test: Ask Questions

- **Plan ahead**
  - Make checklist of things you want to know
- **Ask open-ended questions** for more detailed and accurate information
  - Bad: "Do you understand what this means?"
  - Good: "What do you think when you see this?"
  - Bad: "Did you know you can click here to achieve that?"
  - Good: "What would you do if you would want to achieve that?"
- **Don't blame the user**
  - Bad: "Why didn't you understand this?"
  - Good: "Can you tell me what this means for you?"

# Usability Test: Ask Questions

- **Don't ask the user for solutions**, you are the designer
  - Bad: "Do you need a News button here?"
  - Good: "Which information do you need at this point?"
  - Instead, ask them for comparisons with other systems they've used
- Involving groups (workshops)
  - Users can get into discussions about it, feed off of other opinions
  - Your role: facilitate discussion, avoid intimidation, etc.

# Usability Test: Data Analysis

## Qualitative Data

- Collected from interviews, some types of questionnaires, observation notes
- Interpreted & used for telling a 'story' about what was observed

## Quantitative Data

- Collected from interaction & video logs, closed questionnaires
- Presented as values, tables, charts, graphs and treated statistically

# Usability Test: Findings

- Report on **times to complete task**, **number of errors**
- Provide simple statistical measures (such as for questionnaires): **mean, median, standard deviation**.
- Describe **interaction patterns** observed
  - For example, four ways that people may use the interface

# Usability Test: Relating to Design

- Rank issues **in terms of severity**
- Not only a list of problems and issues!
  - Provide **suggestions on how to address**
- Provide **evidence** (video, quotes, examples) of people encountering issues
- Iterate on the design based on the findings

# Usability Test: Examples

## SpiroSmart & Interactive Data Vis

# Usability Test: Summary

- Users are brought into controlled environments to complete focused tasks
- Focus of the testing is on having representative users and tasks
- Typically, **both objective** (e.g., task completion times, errors) and **subjective data** is collected (e.g., questionnaire data)
- Important end goal: draw meaningful conclusions about your system's current strengths and limitations

# Week 12 Overview

- **Tuesday**

- ~~Introduction to Evaluation~~
- ~~Evaluation: Performance Modeling~~

- **Thursday**

- ~~Evaluation: Inspection~~
- ~~Evaluation: Heuristics~~

- **Friday**

- ~~Evaluation with Users~~