CSE 165: 3D User Interaction

Lecture #17:

UI Evaluation

Videos

- Mitchell Topete: VR Used to treat mental health problems
    - https://www.youtube.com/watch?v=5KIa3nNmMAc
    - https://www.youtube.com/watch?v=2F4XIQWL75c

- Glynn Williams: Tabletop Simulator VR
    - https://www.youtube.com/watch?v=KuUhWPUDpGo

- Shravan Grandhi: Dexmo Exoskeleton
    - https://www.youtube.com/watch?v=Sif7cY8qwjM

Announcements

- Tutors wanted for CSE 190 in spring
- Midterm exam this Thursday
  - Should know what major interaction strategies and devices are called
  - No coding required, but pseudocode possible
  - No Unity or Unreal specific knowledge will be tested
  - No quaternions

Punchkeyboard

- http://l.goodbits.io/l/zr0n6oh2

New Samsung VR Controller

- http://www.roadtovr.com/new-samsung-gear-vr-with-touch-controller-announced-mwc-2017/

Final Project

- Documentation (10 points)
- Presentation (90 points)
    - Technical quality: 70% (10% per skill point)
    - UI Usability: 15%
    - Creativity: 15%
- Extra Credit (10 points)

Due Dates

- Blog Week 1: Sunday, March 12th at 11:59pm
- Blog Week 2: Sunday, March 19th at 11:59pm
- Video: Tuesday, March 21st at 11:30am
- Project Demos: Tuesday, March 21st, 11:30am-2:30pm

Sample Projects from 2016

https://www.youtube.com/results?q=cse+165+2016&sp=EgIQAQ%253D%253D

3D UI Evaluation

Why User Evaluation?

- Need to compare
  - devices
  - interaction techniques
  - Applications
- Problem identification and redesign
- General usability understanding

Some Terminology

- Usability – everything about an artifact and what affects a person's use of an artifact
- Evaluator – person who designs, administers, implements, or analyzes an evaluation
- Subject – person who takes part in the evaluation

Evaluation Tools

- User task analysis
  - generates list of detailed task descriptions, sequences, user work, and information flow
- Scenarios
  - built from task analysis
  - important for experiment design
- Taxonomy
  - science of classification
  - break down techniques into components
  - used in evaluation process
- Prototyping
  - need to have something to test
  - paper-based sketches
  - Wizard of Oz approach

Evaluation Methods

- Cognitive walkthrough
  - Walk through every step of the study with expert
  - Reveals potential usability problems
- Heuristic evaluation
  - Walk through an evaluation checking whether guidelines are being followed
  - Example: check for guideline "Eliminate extraneous degrees of freedom for a manipulation task"
- Formative evaluation
  - observational user studies
  - questionnaires, interviews
- Summative evaluation: compare various techniques in a single experiment
  - task-based usability evaluation: more structured, aimed at improving the interface
  - formal experimentation: have a formal design, analyzed statistically
- Questionnaires
- Interviews and Demos

Evaluation Metrics – System Performance

- System performance metrics
- Average frame rate (fps)
- Average latency / lag (milliseconds)
- Variability in frame rate / lag
- Network delay
- Distortion

- Only important for its effects on user performance / preference
  - frame rate affects presence
  - network delay affects collaboration

Evaluation Metrics – Task Performance

- Speed / efficiency
- Accuracy
- Domain-specific metrics
  - education: learning
  - training: spatial awareness
  - design: expressiveness

Evaluation Metrics – User Preference

- Ease of use / learning
- Presence
- User comfort
- Usually subjective (measured in questionnaires, interviews)

User Comfort

- Simulator sickness
  - Kennedy - Simulator Sickness Questionnaire (SSQ)
- Aftereffects of VE exposure
  - Stanney 1998: Aftereffects from virtual environment exposure: How long do they last?
- Arm/hand strain
- Eye strain

Formality of Evaluation

- Formal
  - independent & dependent variables
  - statistical analysis
  - strict adherence to a procedure
  - hold constant all other variables
  - usually done to compare multiple techniques or at the end of the design process
- Informal
  - looser procedure
  - often more qualitative
  - subject comments very important
  - looking for broad usability issues
  - usually done during the design process to inform redesign

What is Being Evaluated?

- Application:
  - Prototype - consider fidelity, scope, form
  - Complete working system
  - Controlled experiments are rare
- Interaction techniques / UI metaphors
  - Can still evaluate a prototype
  - More generic context of use
  - Formal experiments more often used
- Consider "Wizard of Oz" evaluation

Subjects / Participants

- How many people?
- What backgrounds?
  - technical vs. non-technical
  - expert vs. novice VE users
  - domain experts vs. general population
- What age range?
- Recruiting
  - flyers
  - email/listservs/newsgroups
  - psychology dept.
  - CS classes

Number of Evaluators

- Multiple evaluators often needed for 3DUI evaluations
- Roles
  - cable wrangler
  - software controller
  - note taker
  - timer
  - behavior observer

Procedure

- Welcome
- Informed consent
- Demographic/background questionnaire
- Pre-testing
- Familiarize with equipment
- Exploration time with interface
- Tasks
- Questionnaires / post-testing
- Interviews

Pilot Testing

- Pilot testing should be used to:
  - "debug" your procedure
  - identify variables that can be dropped from the experiment

Instructions

- How much to tell the subject about purpose of experiment?
- How much to tell the subject about how to use the interface?
- Always tell the subject what they should try to optimize in their behavior.
- If using think-aloud protocol, you will have to remind them many times.
- If using trackers, you will have to help users "learn" to move their heads, feet, and bodies – it doesn't come naturally to many people.
- Remind subjects you are NOT testing THEM, but the interface.

Independent Variables

- Main variable of interest (e.g. interaction technique)
- Manipulated by the experimenter
- Conditions under which the tasks are performed
- The number of different values used is called level
    - Example: background can be blue, green, or white (3 levels)
- Secondary variables
    - task characteristics
    - environment characteristics
    - system characteristics
    - user characteristics

Dependent Variables

- Affected by the independent variables
- **Measured** in the user study
- Objective values: e.g., time to complete a task, number of errors, etc.
- Subjective values: ease of use, preferred option
- They should only depend on the independent variables (conditions)

Formal Experiment Issues

- Choosing independent variables
- Choosing dependent variables
- Controlling (holding constant) other variables
- Within- vs. between-subjects design
- Counterbalancing order of conditions
- Full factorial or partial designs

Metrics (dependent variables)

- Task performance time
- Task errors
- User comfort (subjective ratings)
- Observations of behavior (e.g. strategies)
- Spoken subject comments (e.g. preferences)
- Surveys/questionnaires
- Interviews

Data Analysis

- Averages (means) of quantitative metrics
- Counts of errors, behaviors
- Correlate data to demographics
- Analysis of variance (ANOVA)
- Post Hoc analysis (t-tests)
- Visual analysis of trends (esp. learning)

- *Interactions between variables* are often important
- Expect high variance in 3DUI interaction studies
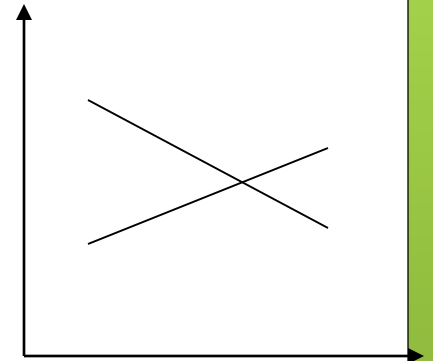
Analysis Tools

- SPSS, SAS, etc.
  - full statistical analysis packages
  - parametric and non-parametric tests
  - test correction mechanisms (e.g., Bonferroni)
- Excel
  - basic aggregation of data
  - Correlations
  - confidence intervals
  - graphs
- Matlab, Mathematica