



## Grading

- Team Assignment Grade (25%)
   T1-T7 (equal contribution)
- Individual Assignment Grade (30%)
   I1-I4 (equal contribution)
- Final Interface/Presentation/Report (25%)
  - T8-T10 (equal contribution) = 20%
  - Peer evaluation = 5%





- April 18, 6pm
  - Email PPT file to Prof & TA
  - Bring on flash drive to class
- Final project presentation
  - Must use template \*\* NOTE UPDATED VERSION \*\*
  - Each presentation will be followed by a few minutes of Q&A
  - Pecha Kucha style
  - Be ready for questions
  - Practice!



# T10

- Refer to instructions for full details
- Introduction
  - What health-related problem were you trying to solve?
  - Why did you think a social application might potentially address this problem?
  - Target users
  - Proper inline citations



# T10

### • Evaluation Method

- How you conducted your user test
  - Evaluation plan that you developed for T8
  - # of users & a basic description of them
     justify that they are from your target demographic
- recruitment approach (how you found users)
- Analysis process (paragraph)
  - Include a photo of your affinity diagram.



# T10

### • Reflection

- What you learned during the iterative design process
- What would you do differently?
  - Don't focus here on design decisions
  - Focus on the meta-level decisions about your design **process**

#### E.g., development of user requirements, ethnographic data collection, facilitating user test sessions, analysis process, etc.



# NEXT WEEK

- Guest Speakers from Google
  - Laura Cuozzo Guarnotta, Senior User Experience Researcher
  - Matt McKeon, Software Engineer
- Read
  - Rubin & Chisnell Ch 12
  - "Critical & Emotional HCI" papers



### **EVALUATION**

- Test & evaluate
  - Functionality
  - Usability
  - Experience
  - Adoption (short + long-term)
  - Impact
    - e.g., behavior/attitude/physiological/organizational/ social change
- Important in all stages in the design life cycle







## EXPERT EVALUATIONS

- Cognitive Walkthrough
- Heuristic Evaluation
- Models (Ch 12, e.g. GOMS)
- Previous work



### WHO ARE WE STUDYING?

- "Computer science [students]: they are simply not representative of the intended user population"
- It is not enough to test with your friends & co-workers

## SAMPLE SIZES

- The number of subjects (participants) in your study
- Nielsen and Landauer
  - One person (1/3 problems)
  - Little to be gained from 5+\*\*
  - Conduct many smaller tests versus one large test
    More cost effective
- Book: recommends at least 10
  - Recent papers ... Even more
- More when testing effects of system
  - Behavior change
  - Group dynamics

## **PILOT TESTING**

### • Pilot testing is critical

- Work out methodological issues
  - Are surveys asking the right question?
  - Does prototype include all functionality you want to test?
  - Is prototype robust enough?
  - Are you getting the kind of data you expect? – E.g., will data help you answer your research questions?
- Ensure testing with right users
- Practice delivery





• Helps you articulate

- Goal of test: what do you want to achieve?

- Different types of goals
  - Quant: "Test", "Prove", "Determine", "Does x impact y"
  - Qual: "Explore", "Examine", "Unpack"



# **TEST PLANS**

• Helps you articulate

- Nature of experimenter-subject interaction
  - Help allowed?
- Data to be collected
- Analysis process
- Criteria for determining the interface is a success



# EVALUATION

What questions do we want to answer about the system?

### **EVALUATION**

- Quantitative data
  - Controlled Experiments
  - Benchmark Testing
  - Physiological Monitoring
- Quant or Qual
  - Naturalistic Studies (deployments)
  - Technology Probes
  - Experience Sampling (can be qualitative)

- Observation is limited
   Misses decision processes and attitudes
- Simple: ask users to speak their thoughts aloud
- Alternative: cooperative evaluation
  - User collaborator in evaluation
  - Work together to identify issues, suggest fixes



#### • Tasks

- Give users a **clear** goal, not the steps
- Start with a simple task
- Realistic
  - Provide a scenario
- Coverage
  - Represent as many possible uses of the system as possible
  - Important parts of interface
- Scope
  - Complete-able in time frame
  - But not trivial



### • Disadvantages

- Can feel unnatural & distracting
- Prematurely prevent errors
- Exhausting

- Demonstrate technique first
- Don't force if strong resistance

   Instead ask probing questions
- Be attentive!
  - User quietness (can signal concentration on problem)
- Show you are listening to them
  - Repeat comments & follow-up



### • Probing

- Ask neutral questions
  - "How did you like the interface?"
  - "Was it hard to understand what that button is for?"
  - YES: "Tell me what you thought about the interface"
  - YES: "Tell me what you thought that button was for. Tell me about how you came to that determination."
    - You can then probe based on their answer. "Oh, it sounds like you had a bit of a challenge, tell me more."
- Ask longer questions in the debrief

## THINK ALOUD

• Study Rubin & Chisnell reading!

# **GUIDELINES FOR USER TESTS**

- You can affect the testing process
  - What you note
  - Verbal & non-verbal communication
  - Takes a great deal of practice to be come skilled
- Moderate Impartially
  - Present product neutrally
  - Don't indicate your approval/disapproval of user comments

# **GUIDELINES FOR USER TESTS**

• Encourage participants to reflect on how <u>they</u> would use system, not others



- Collect multiple kinds of data
- Use qualitative data to expound on quantitative

   Ask "why"?
- Use quantitative data to examine generalizability of qualitative findings
- Helps substantiate your claims



- Contextual Design
  - UCD method
    - Collecting, interpreting & synthesizing data
    - Ethnographically-based
- Affinity Diagramming
  - Technique widely used in HCI
  - Inductive (bottom up)
  - Iterative

### AFFINITY DIAGRAMMING

- Goal
  - Synthesize data
  - Characterize themes, issues, needs etc.
- Analyze qualitative (written) data
  - Interviews
  - Focus groups
  - Observations
  - Task analysis
  - Etc.

are at the cure is an ent are are the technic the t	These are display that are actually run-interactive bot interactive be interactive	Affordances for interaction attract attripted Displays that At social distances of new interactions of new interactions of new interactions interactions interactions Displays that interactions Displays that Displays that Displ	PRUST Storman Water Kolay Mark Like Light	
at the my wolk, t lheight Gus me-ces t hallway	Diring a ges Subway diplay tall, affordstouch but just arv in a window. Rhoro 37 ess Szennigy rominiscour diseaj contains intercedire Junants	Michael de adread Michael de ad	A file can not File for a set one is set one is set enter Pierros Static Pierros Static Pierros Static Pierros Static Pierros Static Pierros P	American     Ameri
er opulated) roll? ge inte	(talat) <b>4</b>	Care PHOTOSE2 NO SEC MONU POOPLOID NOT SDDP 100K Ab brank under Abland Ablan	mentione and the series and the series se	And and any and a set of the set



### • Starting the affinity

- Divide affinity notes amongst team
- Start a column
  - One team member reads & places an affinity note on the wall
  - Others look for related notes & place in column
  - Holtzblatt:
    - 300-400 affinity notes: 1-3/column (not too many 1-note)
    - 500-1000 notes: 4-6
    - Why not many more in a column?
  - If no more related notes found, start a new column

![](_page_23_Figure_12.jpeg)

- Create temporary Green Labels
  - Look at groupings
  - Place above relevant blue labels
  - These labels reflect broad categories
  - "communication strategies"
  - Shoot for 4-6 of these
- Add Pink Labels
  - These are an abstraction of blue labels
  - Each pink label is comprised of 2-6 Blue

![](_page_24_Figure_11.jpeg)

![](_page_25_Figure_1.jpeg)

![](_page_25_Figure_2.jpeg)

• Remember

- Blue & Pink labels
- In the voice of the user, "I…"
- Green labels
  - Categorical
- This is a collaborative effort
  - No one "owns" sections of the wall
- There is no single "right" affinity
  - Validity based on
    - rigor of your method & process
    - qualifications of your team

![](_page_26_Figure_13.jpeg)

# ETHICS

- Researcher's responsibility
  - Help users feel comfortable
  - Don't chastise, make fun of, intimidate, exploit
  - Reinforce their value
  - Convey your gratitude
- Confidentiality
  - Data
  - Participation

![](_page_27_Figure_10.jpeg)

- 1. For this exercise, pick one of your requirements to assess
- 2. Develop strategy for studying how well your system is meeting this requirement
  - Develop a non-trivial task scenario that you will present to users
  - Identify a metric for success (be specific; think trends) e.g., "2/3 of users completed task w/o errors", "All users indicated they would use this feature in the future"
  - ✓ Create a short list of questions that you are hoping to answer through this evaluation & in debrief

*help you elaborate on metrics. E.g., "What aspects of this feature do users like most/least? How could this feature better meet needs?* 

### PAPER PRESENTATIONS

- Wu et al., Tangible Navigation and Object Manipulation in Virtual Environments, TEI'11 Chaitali Kumar (makeup)
- Bernstein et al., Soylent: A Word Processor with a Crowd Inside, UIST 2010 Yucheng Huang
- Gilbert, Designing Social Translucence Over Social Networks, CHI'12 Herman Saksono
- Srinivivasan, Ethnomethodological Architectures: Information Systems Driven by Cultural and Community Visions, 2007 Luobin Bai