CSE 510: Advanced Topics in HCI

HCI as Design II

James Fogarty
Daniel Epstein

Tuesday/Thursday
10:30 to 12:00
CSE 403
I used to think correlation implied causation.

Then I took a statistics class. Now I don't.

Sounds like the class helped.

Well, maybe.
Reporting

“extremely” or “very” significant wording issue
significance vs. effect size

“slightly” or “barely” significant wording issue
significance vs. effect size
“marginally significant” for $p < .10$, also a “trend”

“insignificant” is not a term
no significant difference
not able to detect a significant difference
Reporting

Communities have reporting norms

\[ p < 1 \times 10^{-7} \text{ vs. } p < 0.0001 \]

Test statistics often reported at 3 digits
f test has Between and Within DOF, often rounded

Provide higher-level takeaways
not just a wall of stats
meaningful conditions names when possible
careful with abbreviations
qualitative content can complement

Careful in wording of claims vs. results of tests
"Interface C also leads to fewer restarts than interface B, but we cannot determine that the difference is significant".
Roles of Variables

Fixed vs. Random effects

If you ran the experiment again, would you have the same values for this variable?

Fixed: “Data has been gathered from all the levels of the factor that are of interest.”

Random: “The factor has many possible levels, interest is in all possible levels, but only a random sample of levels is included in the data.”

Know whether you are including a variable as a control or for an experimental outcome e.g., analyzing task
Feature Selection in Models

Correlated factors might improve model fit, but might not be what you want to study (i.e., overfitting)
- e.g., CalendarDay and StudyDay are highly correlated
- e.g., effect size may reveal features are offsetting
- Even random variables can be fit to data

Communities have differing norms
- Main effect then pairwise contrasts
- Explain what interactions were tested and why
- Automated feature selection is uncommon in CHI

Stats are fundamentally a tool for hypothesis testing
- Extreme interpretation is that you should have your model designed even before you do the study
“Do the Work” vs “Understand It”

HCI practice includes both

CSE 440 teaches an intense project sequence
Interjects higher-level understanding

Today will focus on conceptual material
Thursday will focus on a typical design process

Highly abridged presentation of this material
Learning to Give and Receive Critique

You will learn how to both give and receive critique

Each is important
Each is a skill developed through practice

Many activities will consist of group critiques

Each group will present an artifact
Other class members and staff will offer critique

Starting today with critique of the CI Plan
Why Critique?

Critique helps evaluate early, often, and cheaply

Applicable to artifacts of many types
Compare to other expert inspection methods

You are not your own worst critic

We collectively know more than any one of us
It is hard to see past your own decisions
Design requires getting past our own infatuation

A design can feel like our love, our baby...
Critique is About Improvement

http://alistapart.com/article/design-criticism-creative-process
Tips for Critics: Hamburger Method

“Bun, meat, bun”

Bun:
Something fluffy and nice

Meat:
Criticism on how to improve

Bun:
Something fluffy and nice

Not a “shit sandwich”
Positives need to be genuine, enable learning from both positive and negative aspects of the artifact
Tips for Critics: I Like, I Wish, What If

I Like:
   Lead with something nice

I Wish:
   Some criticism, often leading from what you like

What If:
   An idea to spark further conversation, better than: “I think you should have...” or “Why didn’t you ...”
   Gives the presenter benefit of the doubt if they did already think of your idea, can present rationale
Tips for Critics: Socratic Method

Identify an aspect of the design and ask “Why?”

Can be good if unsure what else to say
Forces presenter to give, or develop, explanations for decisions, which can help build design rationale
Not fundamentally negative and hard to get defensive
“You Are Not the Customer”

Seems obvious, but...

You have different experiences
You have different terminology
You have different ways of looking at the world

Easy to think of self as typical

Easy to make mistaken assumptions
Ethnography

Traditional science attempts to understand a group or individual objectively

Understand the subject of study from the outside in a way that can be explained to “anyone”

Ethnography attempts to understand a group or individual phenomenologically

Understand the subject of study as the subject of study understands itself
Ethnography

Emerged in 1920s as a new anthropology method, exploring why groups think and act as they do

Learn local language, record myths, customs, and ceremonies in much greater detail than prior work

You will likely never perform an ethnography
Four Ethnographic Principles

Natural settings
Holism
Descriptive
Member point-of-view
Four Ethnographic Principles

Natural Settings

Conducted in the setting of the participant

Focus on naturally occurring, everyday action

Cannot use laboratory, experimental settings, or a phone call to gather this type of data

You really do have to go out there and see it
Four Ethnographic Principles

Holism

Behavior can only be understood in its larger social context; that is, holistically.
Four Ethnographic Principles

Descriptive

Study how people actually behave, not how they ought to behave.

Defer judgment.
Four Ethnographic Principles

Member Point-of-View

See through participant eyes in order to grasp how they interpret and act in their world.
Contextual Inquiry

Applied design ethnography

“The core premise of Contextual Inquiry is very simple: go where the customer works, observe the customer as he or she works, and talk to the customer about the work. Do that, and you can’t help but gain a better understanding of your customer.”

Hugh Beyer and Karen Holtzblatt
What is your relationship?

In a scientist/subject relationship:

The scientist does stuff
The subject responds in some way
The scientist collects data, goes back to their office, and analyzes the data to gain understanding

This is not very appropriate for gaining phenomenological understanding
User, Subject, or Participant?

Only two groups refer to their customers as users.

In traditional science, “subjects” are “subjected to” experiments as a researcher develops understanding.

In ethnographically-oriented design methods, “participants” instead “participate” in helping the researcher develop understanding.

This isn’t simple PC, it’s a mindset that matters.
What is your relationship?

In an interviewer/interviewee relationship:

- The interviewer asks a question
- The interviewee responds immediately
- At a pause, the interviewer asks another question from a list
- When all the questions are answered, the interview is over

This would only be appropriate for gaining phenomenological understanding if you knew what questions to ask in advance

Implying you have phenomenological understanding
What is your relationship?

In a master/apprentice relationship:
- The master is doing stuff
- The master explains what they are doing
- The apprentice asks clarification questions
- The master answers

This relationship is at the heart of contextual inquiry
Master/Apprentice Relationship

Seeing the work reveals structure

Many instances and many interviews reveal the picture

Every current activity recalls past instances

A customer describing how she learned a feature told us, “I looked it up in the documentation.” But when we asked her to look it up again, she was able to show us: “I looked the function up in the index and scanned the section. I saw this icon in the margin that I recognized from the screen, so I read just this paragraph next to it. It told me all I needed to know.” The documentation provided the context she needed to recover a detailed story, and the detail revealed aspects that had been overlooked—that the icon was her visual cue to the relevant part of the page.
Not Quite Master/Apprentice

In a contextual inquiry relationship:
- The participant is doing stuff
- The participant explains what they are doing
- The researcher offers an interpretation
- The participant agrees or corrects

Partners
- Not really an interview
- Not really an apprentice
Principles of Contextual Inquiry

Context
  Must be done in the setting of the participant.

Partnership
  Master/apprentice model; investigator is humble.

Interpretation
 Observed facts must be regarded for their design implications. Raw facts without interpretation are not very useful.

Focus
  Themes that emerge during the inquiry. You cannot pay attention to all facets of someone’s work at all times.
Context

Go to the workplace & see the work as it unfolds
People summarize, but we want details
Keep it concrete when people start to abstract
“Do you have one? May I see it?”
Context

Imagine studying how a student writes a paper

Why not just ask?

May not remember details

- Getting roommate to read drafts

May skip critical difficulties

- Trouble locating references on the Web
Context

Avoid summary data by watching work unfold

We once asked a secretary how she started her day. Her answer was, “I guess I just come in and check my messages and get started.” She wasn’t able to go beyond this brief summary overview. It was the first thing in the morning and she had just arrived at the office, so we asked her to go ahead and do as she would any other morning. She unhesitatingly started her morning routine, telling us about it as she went: “First I hang up my coat, then I start my computer. Actually, even before that I’ll see if my boss has left something on my chair. If he has, that’s first priority. While the computer’s coming up, I check the answering machine for urgent messages. There aren’t any. Then I look to see if there’s a fax that has to be handled right away. Nope, none today. If there were, I’d take it right in and put it on the desk of whoever was responsible. Then I go in the back room and start coffee. Now I’ll check the counters on the copier and postage meter. I’m only doing that because today’s the first of the month. . . .”

Have them think aloud..
Partnership

Traditionally, interviewer has too much power
You don’t know what will turn out to be important

Apprenticeship model tilts power back too far
You aren’t there to learn the skill

Interviewer should create a partnership
Alternate between watching and probing
Partnership

Withdrawal and return

Researcher observes action that indicates something meaningful

The researcher asks about this, and the pair withdraw from the task

Discuss the question

Then return to the task

In one interview with a user of page layout software, the user was positioning text on the page, entering the text and moving it around. Then he created a box around a line of text, moved it down until the top of the box butted the bottom of the line of text, and moved another line of text up until it butted the bottom of the box. Then he deleted the box.

Interviewer: Could I see that again?
Customer: What?
I: What you just did with the box.
C: Oh, I'm just using it to position this text here. The box doesn't matter.
I: But why are you using a box?
C: See, I want the white space to be exactly the same height as a line of text. So I draw the box to get the height. (He repeats the actions to illustrate, going more slowly.) Then I drag it down, and it shows where the next line of text should go.
I: Why do you want to get the spacing exact?
C: It's to make the appearance of the page more even. You want all the lines to have some regular relationship to the other things on the page.
Interpretation

Chain of Reasoning

Fact, Hypothesis, Implication for Design, Design Idea

Design is built upon interpretation of facts

Design ideas are end products of a chain of reasoning

So interpretation had better be right

Share interpretations with users to validate

Will not bias the data

Teaches participant to see structure in the work
Interpretation

Instead of asking open ended questions...

“Do you have a strategy to start the day?”
“Not particularly.”

... give participants a starting point

“Do you check urgent messages first, no matter where they are from?
“Actually, things from my boss are important, because they are for me to do. Messages or faxes may be for anybody.”

Participants fine-tune interpretations

Probe contradictions until assumptions fit
Focus

Everybody has a focus, you cannot prevent it

Entering focus
Project focus

Because you will have a focus, be mindful of that focus and use it to your advantage

Brainstorm and define your focus
The Stages of a Contextual Inquiry

1. Interview / Warm Up
2. Transition
3. Observe Behavior
4. Share Interpretation
5. Refine Interpretation
6. Wrap-up
Affinity Diagrams

Generated during group session

Each observation, idea, note to a post-it

Notes are hierarchically organized into themes, based on project focus
Flow Model: Secretarial Hub

President
- Run the business
- Keep abreast of what's going on
- Sign checks
- Go on trips

Worker
- Do the work of the business
- Meet with management

Vendor
- Invoice for services

Sales manager
- Run the sales department
- Travel to sales offices

Marketing manager
- Run the marketing department
- Produce proposal

Secretary
- Keep office organized
- Ensure bills paid on time
- Do final proof, print, and distribution of documents
- Manage and coordinate schedules
- Handle logistics of trips

Department's reports
- Checks to sign

Request to help with family vacation plans

Request to schedule meeting with president

Invoices
- Checks

Proposal to proof and mail

Discussion of travel plans

Announcement

Bulletin board
- Announce events of general interest
- Hold documents that manage shared projects

U1

Requires lots of iterations
Sequence Model: Doing Email

Intent: Handle emergencies

- Trigger: Return to the office
- Scan message list for important message—Use sender, subject
- Choose urgent message
- Read message about unhappy user
- Decide more info needed
- Make phone call
- Had to put off issue of unhappy user

Intent: Get back to people easily

- Leave phone message
- File in phone folder
- See list of messages
- Choose message 9: subject indicates university news relevant to department
- Read message
- Delete message
- See message 10 automatically
- Read message 10
**Sequence Model: Equipment Audit**

1. Assigned to do equipment audit
2. Retrieve required form from database
3. Print form
4. Collect data at site
5. Record data on paper form
6. Type data into form on computer
7. Print completed form
8. Leave hardcopy of form with customer
9. Send electronic form to supervisor
10. Store electronic form on form database
Cultural Model: Developer

- Marketing
  - Our new features are top priority
  - If I say do X, you figure out what that means

- Competitors
  - We have 50 new features; catch up

- Base technology group
  - You aren't our primary user; we'll fix bugs for you in our own time
  - Our technology is standard; use it even if it doesn't work

- Customer support
  - Our bug reports are top priority

- U9 (Developer)
Artifact Model: Calendar

Past (*seldom accessed*) - Future (*quick access*)

- **Scheduled events**
- **Unscheduled but associated with the day**
- **Reminders (storage with quick access)**

- **Meetings**
- **Appointments**
- **Reminders**
  - Strike out a day
  - Notes
  - Never used

Business cards (*storage for later*)

Rubber band
Physical Model: Work Site

- Maybe outside
- Large area (up to square mile)
- Tight spaces
- Climbing
- Awkward positions

Company Trailer

Computer

Approximately a 5 minute walk. If doing an audit at a site under construction, then safe path frequently changes and may need to wait for construction equipment to pass.
Tasks Matter

System will fail if:

  It is inappropriate for the customer
  It does not meet customer needs

Your contextual inquiries will emphasize getting to know your customers and their needs

Can’t you then just make ‘good’ interfaces?
Why Task Analysis?

‘Good’ has to be interpreted in the context of use
   Might be acceptable for office work, but not for play
   Infinite variety of tasks and customers

Guidelines are too vague to be generative
   e.g., “give adequate feedback”
   Can be used to critique, but not to generate

Design is often about tradeoffs
Why Task Analysis?

Task analysis is a lens on the information you obtain through methods like contextual inquiry. Use what you learned in your inquiry to answer the questions in the task analysis.

Your assignments order the two, but in practice you should iteratively decide how to best draw upon all relevant methods throughout a process.
11 Task Analysis Questions

Who is going to use the system?
What tasks do they now perform?
What tasks are desired?
How are the tasks learned?
Where are the tasks performed?
What is the relationship between people & data?
What other tools do people have?
How do people communicate with each other?
How often are the tasks performed?
What are the time constraints on the tasks?
What happens when things go wrong?
Selecting Tasks

Real tasks people have faced or requested
collect any necessary materials
Should provide reasonable coverage
compare check list of functions to tasks
Mixture of simple and complex tasks
easy tasks (common or introductory)
moderate tasks
difficult tasks (infrequent or for power use)
What Should Tasks Look Like?

Say what person wants to do, but not how

allows comparing different design alternatives

Be specific, stories based in concrete facts

say who person is (e.g., using personas or profiles)

design can really differ depending on who
give names (allows referring back with more info later)
characteristics of person (e.g., job, expertise)

story forces us to fill in description with relevant details

Sometimes describe a complete “accomplishment”

forces us to consider how features work together
Using Tasks in Design

Write up a description of tasks formally or informally run by people and rest of the design team get more information where needed

Manny is in the city at a restaurant and would like to call his friend Sherry to see when she will be arriving. She called from a friend’s house while he was in the bus tunnel, so he missed her call. He would like to check his missed calls and find the number to call her back.
Task: Park in a New Neighborhood

Peter is going to brunch on a Sunday with his roommates. He is trying a new place he found on Yelp. He has the address for the place and he is using a smartphone GPS for directions. He leaves the apartment with his roommates at around 8:30am and he wants to beat the crowd so they won’t have to wait in line. He is driving a Toyota Corolla that he has owned for five years. It is a rainy day and he doesn’t have an umbrella.
Hierarchical Task Analysis

Steps of the task execution (detailed in a hierarchy)

park in new neighborhood

- determine destination
- drive to destination
- locate parking spot
- secure parking spot
- park

- enter address in GPS
- follow directions
- arrive at destination

...
Hierarchical Task Analysis

Steps of the task execution (detailed in a hierarchy)

park in new neighborhood

- determine destination
- drive to destination
- locate parking spot
- secure parking spot
- park

- enter address in GPS
- follow directions
- arrive at destination

...
Using Tasks in Design

Rough out an interface design
discard features that do not support your tasks
or add a real task that exercises that feature
major elements and functions, not too detailed
hand sketched

Produce scenarios for each task
what person does and what they see
step-by-step performance of task
illustrate using storyboards
Scenarios

Scenarios are design specific, tasks are not
Scenarios force us to
show how things work together
settle arguments with examples
but these are only examples, and
sometimes need to look beyond flaws

Show people storyboards
nice mechanism for feedback
Tasks, Personas, and Scenarios

**Task**: a design-agnostic objective

**Persona**: a fictional person with a backstory

**Scenario**: narrative that demonstrates a persona completing a task using a particular design

**Use Case**: in software engineering, describes requirements using one or more scenarios
Sketching and Storyboards

<table>
<thead>
<tr>
<th>Attendance List</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee, Benjamin</td>
</tr>
<tr>
<td>Santos, Allen</td>
</tr>
<tr>
<td>Schwartz, Jonah</td>
</tr>
<tr>
<td>Vermette, Joshua</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sort By</th>
<th>Show</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last Name</td>
<td>Enroll</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Last Name</th>
<th>First Name</th>
<th>SID #</th>
<th>Enrollment</th>
<th>Section</th>
<th>Major</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwartz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vermette</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Enroll</th>
<th>Waithlist</th>
<th>Audit</th>
<th>Presence</th>
<th>Absent</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Go to Attendance View

Back to main menu

38 Present, 2 Absent

Take Attendance from students' PDA

Done

Look Up: So... highlights student
Sketching and Storyboards

SCENARIO 1: "I want to listen to alternative music"

Diagrams showing navigation between different music categories such as Top 40, Alternative, Country, Metal, R&B, and Soul.
Sketching and Storyboards
Illustrating Time

Storyboards come from film and animation

Give a “script” of important events
leave out the details
concentrate on the important interactions
Storyboards

Can illustrate key requirements and leave open less important details of design
Basic Storyboard

CREATE MY RIDE PROFILE
RESEARCH ROUTE
SAVE ROUTE TO MY RIDE
RECEIVE UPDATE

ARRIVE AT STOP
ON BUS
BUS DISPLAY
PROVIDE FEEDBACK

MyRide
Sign Up
Name
Email
Mobile
Home bus stop's

From: Home
To: Whole Foods
Options
6 bus
Time
12:14

MyRide:
Home to Whole Foods
Leave: Sat, 8am
Return: Sat, noon
Route: 7 bus

Nokia
MESSAGES
MYRIDE:
5 min delay on 7 bus.
ETA at bus stop 7:05am

On Bus

The Ride
Next stop: Stadium
12:21
Stop
Stop
Stop

Arrive at Stop

Bus Display

Provide Feedback

Open Letter:
From: MyRide
Subject: Feedback

Thanks for taking "7 bus to WholeFoods" How was your ride? Click for feedback
Storytelling

Stories have an audience

Other designers, clients, stakeholders, managers, funding agencies, potential end-users

Stories have a purpose

Gather and share information about people, tasks, goals
Put a human face on analytic data
Spark new design concepts and encourage innovation
Share ideas and create a sense of history and purpose
Giving insight into people who are not like us
Persuade others of the value of contribution
Stories Provide Context

Characters
- Who is involved

Setting
- Environment

Sequence
- What task is illustrated
- What leads a person to use a design
- What steps are involved

Satisfaction
- What is the motivation
- What is the end result
- What need is satisfied

Details of interface features and components are not necessarily surfaced, they can often be developed and conveyed more effectively with other methods.

Can help surface details that might otherwise be ignored.

Grocery store application:
- use with one hand while pushing a shopping cart
- privacy of speech input
- split attention
Elements of a Storyboard

Visual storytelling

5 visual elements

- Level of detail
- Inclusion of text
- Inclusion of people and emotions
- Number of frames
- Portrayal of time

To better characterize design intuitions: gather and analyze artifacts, semi-structured interviews, survey focused on identified elements

Truong et al, 2006
1. How Much Detail?
1. How Much Detail?

Unnecessary details distract from the story.
2. Use of Text

Guideline: It is often necessary, but keep it short

1. At home, Mary checks her blood pressure.
2. After a few simple key presses, her blood pressure readings get sent to a clinic.
3. The information is made available to her doctor.

Short text is more effective, less likely to over-explain.
Watch for cases where text induces weird biases.
3. Include People and Emotions

Guideline: Include people experiencing the design and their reactions to it (good or bad)

Remember, the point of storyboards is to convey the experience of using the system
4. How Many Frames?

Guideline: 4-6 frames is ideal for end-users

- Less work to illustrate
- Must be able to succinctly tell story
- Potentially longer for design clients

More is not always better

- May lose focus of story
- May lose attention
5. Passage of Time

Guideline: Only use if necessary to understand

Inclusion of the clock distracts
Storyboards for Comparing Ideas

Cooperative

Let's use our cell phones to keep a record of the number of days that we exercise!

1st Week

2nd Week

Yeah! We are almost there. Good job!

Okay! Let's work together to meet a goal of exercising for least 2 weeks.

Competitive

Let's compete to see who exercises more.

1st Week

2nd Week

Let's see who wins next week.

Okay, Let's do it!

Yeah! I win this week!
Value of Animation or Video

Can illustrate critical timing

Can be more engaging than written or storyboard

Can more easily convey emotion (e.g., voice, music)

Can show interactive elements more clearly

Can be self-explanatory

*If done well, can be an effective pitch*

But you need to keep it quick and effective
Prototyping Microsoft Surface

Prototyping Microsoft Surface

Split Presentation, Simple Effects

http://courses.cs.washington.edu/courses/cse440/videos/videoprototyping/Pickup.mp4
Sun’s “Starfire” (1994)

http://courses.cs.washington.edu/courses/cse440/videos/videoprototyping/Vision-Sun-Starfire.mp4
Apple’s “Knowledge Navigator” (1987)

Corning’s “A Day Made of Glass” (2011)

Is My Design Good?

This is not a meaningful question

   It can and will be answered with “Yes”

At least consider asking:

   “What are three good things about this design?”
   “What are three bad things about this design?”

But really the answer is “it depends”

Remember that designs are used for tasks
We should ask this in the context of tasks
Paper Prototype

<table>
<thead>
<tr>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td>8-9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Welcome to ESP.
Your Telebeats session is Tues. Sept. 21 @ 10am
Your current schedule is empty. Please click on Add a course to continue.

“Screen” faked with pre-constructed pieces
Paper Prototype

New pieces added in response to interaction
Paper Prototype

Transparencies allow flexible use of text
Paper Prototype as Communication

1. Personal Details
2. Address
3. Account
4. Trading Experience
5. Agreement
6. Finish

Account Type:
Username:
Password:
Confirm Password:
Account Currency: GBP
Comm's Presence: Email using

Account Type
Advanced Account
Basic Account
Paper Prototype as Evaluation

<table>
<thead>
<tr>
<th>Daum</th>
<th>한국어</th>
<th>입력</th>
<th>검색</th>
<th>검색 결과</th>
</tr>
</thead>
<tbody>
<tr>
<td>1234</td>
<td>lorem</td>
<td>ipsum</td>
<td>dolor</td>
<td>sit amet</td>
</tr>
</tbody>
</table>

Table:
- 1234: [Information]
- lorem: [Information]
- ipsum: [Information]
- dolor: [Information]
- sit amet: [Information]
Constructing the Prototype

Remember your target platform constraints
Inspection-Based Methods

We have cut prototyping to its minimum

Sketches, storyboards, paper prototypes
Rapid exploration of potential ideas

But we need evaluation to guide improvement

Evaluation can become relatively slow and expensive
Study participants can be scarce
May waste participants on fairly obvious problems
Inspection-Based Methods

Simulate study participants

Instead of actual study participants, use inspection to quickly and cheaply identify likely problems

Inspection methods are rational, not empirical

Today we cover two complementary methods

Heuristic Evaluation
Cognitive Walkthrough
Heuristic Evaluation

Developed by Jakob Nielsen
Helps find usability problems in a design
Small set of evaluators examine interface
  three to five evaluators
  independently check compliance with principles
  different evaluators will find different problems
  evaluators only communicate afterwards

Can perform on working interfaces or sketches
Nielsen’s 10 Heuristics

Too few unhelpful, too many overwhelming

“Be Good” versus thousands of detailed rules

Nielsen seeks to create a small set

Collects 249 usability problems
Collects 101 usability heuristics
Rates how well each heuristics explains each problem
Factor analysis to identify key heuristics

Nielsen, 1994
Nielsen’s 10 Heuristics

Visibility of system status
Match between system and the real world
User control and freedom
Consistency and standards
Error prevention
Recognition rather than recall
Flexibility and efficiency of use
Aesthetic and minimalist design
Help recognize, diagnose, and recover from errors
Help and documentation

Nielsen, 1994
Phases of Heuristic Evaluation

1) Pre-evaluation training
   give expert evaluators needed domain knowledge & information on the scenario

2) Evaluation
   individuals evaluate interface & make lists of problems

3) Severity rating
   determine how severe each problem is

4) Aggregation
   group meets & aggregates problems (w/ ratings)

5) Debriefing
   discuss the outcome with design team
How to Perform Evaluation

At least two passes for each evaluator
  first to get feel for flow and scope of system
  second to focus on specific elements
If system is walk-up-and-use or evaluators are domain experts, no assistance needed
  otherwise might supply evaluators with scenarios
Each evaluator produces list of problems
  explain why with reference to heuristic
  be specific & list each problem separately
Example Heuristic Violation

1. [H4 Consistency]

The interface used the string "Save" on the first screen for saving the user's file, but used the string "Write file" on the second screen. Users may be confused by this different terminology for the same function.
Severity Rating

Used to allocate resources to fix problems
Estimates of need for more usability efforts
Combination of
  frequency
  impact
  persistence (one time or repeating)
Should be calculated after all evaluations are in
Should be done independently by all judges
Severity Rating

0 - Do not agree this is a problem.

1 - Usability blemish. Mild annoyance or cosmetic problem. Easily avoidable.

2 - Minor usability problem. Annoying, misleading, unclear, confusing. Can be avoided or easily learned. May occur only once.

3 - Major usability problem. Prevents users from completing tasks. Highly confusing or unclear. Difficult to avoid. Likely to occur more than once.

4 - Critical usability problem. Users will not be able to accomplish their goals. Users may quit using system altogether.
Example Heuristic Violation

1. [H4 Consistency] [Severity 3]

The interface used the string "Save" on the first screen for saving the user's file, but used the string "Write file" on the second screen. Users may be confused by this different terminology for the same function.
Fixability Scores

1 - Nearly impossible to fix. Requires massive re-engineering or use of new technology. Solution not known or understood at all.

2 - Difficult to fix. Redesign and re-engineering required. Significant code changes. Solution identifiable but details not fully understood.

3 - Easy to fix. Minimal redesign and straightforward code changes. Solution known and understood.

4 - Trivial to fix. Textual changes and cosmetic changes. Minor code tweaking.
Example Heuristic Violation

1. [H4 Consistency] [Severity 3] [Fix 4]

The interface used the string "Save" on the first screen for saving the user's file, but used the string "Write file" on the second screen. Users may be confused by this different terminology for the same function.

Fix: Change second screen to "Save".
Why Multiple Evaluators?

Every evaluator doesn’t find every problem

Good evaluators find both easy & hard ones
Decreasing Returns

problems found

benefits / cost

Nielsen, 1994
Alternative Inspection-Based Methods

Cognitive Walkthrough

Helps surface different types of usability problems
Consider this as a complement to heuristic evaluation

Action Analysis

Low-level modeling of expert performance
Be aware of GOMS, but you may never encounter it
Cognitive Walkthrough

Evaluation method based on:

- A person works through an interface in an exploratory manner
- A person has goals
- The person is applying means-ends reasoning to work out how to accomplish these goals

Evaluation by an expert, who goes through a task while simulating this cognitive process
Preparation: Need Four Things

1) User description, including level of experience and any assumptions made by the designer
2) System description (e.g., paper prototype)
3) Task description, specifying the task the expert has to carry out, from a user’s point of view
4) Action sequence describing the system display and the user actions needed to complete the given task. One system display and one user action together are one step.
Cognitive Walkthrough Process

Designer/Developer prepares the required documents described on previous slide

Gives these documents to the usability expert

Expert reads the descriptions, and carries out the task by following the action list

At each step in action list, asks four questions

Record problems similar to heuristic evaluation
Believability

1) Will the user be trying to produce whatever effect the action has?
2) Will the user be able to notice that the correct action is available?
3) Once the user finds the correct action at the interface, will they know that it is the right one for the effect they are trying to produce?
4) After the action is taken, will the user understand the feedback given?
Action Analysis / Cognitive Modeling

GOMS: Goals, Operators, Methods, Selection

Developed by Card, Moran and Newell

Walk through sequence of steps
Assign each an approximate time duration
Sum to estimate overall performance time

1. Select sentence
   Reach for mouse H 0.40
   Point to first word P 1.10
   Click button down K 0.60
   Drag to last word P 1.20
   Release K 0.60
   3.90 secs
Inspection vs. Usability Testing

Inspection is

- Is much faster
- Does not require interpreting user actions
- May miss problems or find false positives

Usability testing is

- More accurate, by definition
- Account for actual users and tasks

One approach is to alternate between them

- Find different problems, conserve participants
Deciding What Data to Collect

Process data
- Observations of what people do and think
- Focused on improving this process

Summary, statistical, or bottom-line data
- Summary of what happened (time, errors, success)
- Focused on measurement

Focus on process data
- Gives overview of where the problems are
- More useful than “too slow” or “too many errors”
Not a Scientific Experiment

Focus is on improving the design

- Experimental control is not as necessary
- Data measurement is not as precise
- Number of participants is fairly small

Changes can be made

- Fix the obviously broken design
- Quickly explore alternatives
- Modify the focus of testing between participants
Task-Based Usability

Set up an overall context

“We are interested in improving people’s ability to save, update, and use contacts in their mobile phones.”

Then prescribe tasks

1. Try to find the contacts list in the phone
2. View the contact information for John Smith
3. Change John Smith’s number to be 555-555-5555

Tasks can be chained to naturally lead to the next...
Stages of a Usability Test

Preparation
Introducing the Test
Conducting the Test
Debriefing
Analyzing the Data
Creating the Report
Preparing for a Test

Select your participants
Friends and family are not your design targets
Understand background, consider recruiting questionnaire

Prepare tasks and paper prototype

Practice to avoid “bugs” in your prototype
Usability Test Proposal

A report that contains

Objective, Description of System, Environment and Materials, Participants, Methodology, Tasks, Test Measures

Work through it with colleagues to debug test

Reuse when presenting final report
Introducing the Test

Address Feelings of Judgment

“Today we are interested in learning about X. That’s where you come in!”

“I did not develop X. I just want to know what the problems are with X.”

“It is X being tested here, not you.”
Introducing the Test

Set Expectations for Process

“It is essential you think out loud while working with X. Tell me constantly what you are thinking, looking for, wondering, confused about, surprised, and so on. If you stop talking, I will prompt you to talk.”

“I will not be able to answer your questions when you start using X. Do you have any questions now?”
Conducting a Test

See the Gommol reading tips on a test session
Talk-Aloud Prompts

“Tell me what you are trying to do.”

“Please keep talking.”

“Tell me what you are thinking.”

“Are you looking for something? What?”

“What did you expect to happen just now?”

“What do you mean by that?”

“Talk-aloud” is similar but distinct from “think-aloud”

Most do not know or care about the difference, so you may see the terms used interchangeably.
Insight Problems

When people are trying to figure something out, talking aloud can prevent needed “insight”

If your participant is really baffled, it might not be the best time to prompt them to keep talking

Wait for a natural break, and then ask “What were you thinking just there?”

Retrospective talk-aloud

Record session, talk through immediately afterward
Answering Questions

Remember the purpose of this test

You would not be there “in real life”
You want to see if they can figure it out
You want to see how hard it is
You want to see how catastrophic the outcome is

But you do not want to punish the person or completely undermine the rest of the session

Note any help you provide as a major failure
Do not allow observing engineers to help
Debriefing

Give them more details about what you were interested in discovering, with their help

Answer any questions they have

Now you can show them how to accomplish the tasks, talk about what you learned from the test

Thank them for their time

Appropriate to give some compensation
Analyzing and Reporting the Results

Tests yield many forms of data

Quantitative counts
  time, success/failure
  confusions, errors, workarounds

Observations
  notes about when, where, why, how above occur

Participant comments and feedback
  during session of via a questionnaire
Analyzing and Reporting the Results

Summarize the data

Make a list of critical incidents
- can be positive and negative
- include references back to original data
- try to judge why each difficulty occurred

Sort and prioritize findings
- what does data tell you
- what are the important results
- anything missing from test
Task Design is Important

The goal of a test is to figure out how a person interacts with an interface in the wild...

There are two possible explanations for why a test does not find significant problems:

The interface does not have significant problems

The test itself has significant problems
Task Design is Important

Testing is not entirely in the wild

As a part of focusing the test, you often need to give a person a somewhat artificial task

The artificiality of the task may influence how people interact with an interface...

...and thus may influence the outcomes and insights gained through user testing
Bad: Artificial Subgoals

People using the design “in the wild” may not necessarily form these same subgoals.

The task should give one top-level goal, a people should form their subgoals while pursuing this.

Now you want to choose the type of paper you want to print your document on. Lets imagine that Bin “B” has the paper you want to print your paper on, please complete this task.

Now set the darkness of your copies to about 50% dark. After setting the darkness, you decide you want to print 2 sides of copies on two sides of paper. Please complete this task.
Bad: Artificial Ordering

With an artificial ordering of information or subgoals, people might not proceed in this order.

The ordering might also be biased towards the layout of the interface, which would conceal any problems with finding the appropriate control.

- Enter in 10 copies, with lightness set to 10%.
- Choose 1 sided to 2 sided, use paper source bin A.
- Cover sheet needed, using paper bin B for cover sheet.
- Set stapling feature on and collating on.
- Start printing.
Bad: Changing the Task

The task is to make copies, and this happens to involve entering information in the copier interface.

But this task description is an data entry task, “Here is some information. Put it in the interface.”

- Make 23 copies
- With collate
- Cover sheets
- Default darkness
- 1 Sided-> 1 Sided
Bad: Giving the Answers

Tells the person what terminology the interface uses, which they might not otherwise know

lighten = contrast, sorted = collated?

You are a teacher and are trying to make 40 copies of a one-sided magazine article that is 10 pages long for your class tomorrow. Due to the large number of copies, you print the article double-sided, in other words 10 page article would be printed on 5 sheets of paper. Due to the high contrast of the article, you must lighten the copy, in other words change the contrast. You then want the copies to be collated and stapled.
Good: Giving Context

Giving realistic context through scenarios can reduce the artificiality of the task

It’s your first day in the office, starting a new job. You would like to make some copies of several documents that your boss gave you to browse through. Your colleague in the next cubicle tells you that you need an access code to make copies. The code is 5150. You walk over to the copy machine at the end of the hall and realize that it is not the Xerox copier that you are accustomed too... Make 2 copies of the “Company Annual Report”.

Consider: Under-Specified Tasks

Many realistic goals are under-specified, as people have only a general idea what they want

By under-specifying the task, you can elicit realistic confusion and decision-making

You just finished fixing up the old hot rod in the garage and now its time to sell her. Make a couple copies of the pictures you took to send into the used car sales magazines. It’s ok that they’re in black and white but maybe you should lighten them up a bit. Your account billing code is 5150.
Task Design Summary

Task design is difficult and important

Poorly designed tasks mask interface failures

If you are not confident in your task descriptions, have others help you “debug” them before testing
CSE 510: Advanced Topics in HCI

HCI as Design II

James Fogarty
Daniel Epstein

Tuesday/Thursday
10:30 to 12:00
CSE 403