Today's Agenda

- Evaluation methods
  - Card Sorting
  - Heuristic evaluation
  - Field studies

Individual Assignment

- Written critiques (all students)
- A case study of evaluation (graduate students ONLY)
- Due 11:59 pm EST, Tuesday, Oct. 29th
- Submit via Blackboard

Group Assignment: Evaluation Plan

Each group needs to submit a list of evaluation methods you plan to use

Title: "CSCE 572 Group X Evaluation Methods"

- >= 3 different evaluation methods
- Rationale of using the methods
- Provide some detail on how you plan to implement them, e.g., a protocol
- ~ one page
- Due at 11:59pm, Nov. 12 in Blackboard

Sorting Things Out: Card Sorting Methodology

Open Sort vs. Closed Sort

Open Sort

- Participants are asked to organize topics from content within your website into groups **that make sense to them**
- Then **they name each group** they created in a way that they feel accurately describes the content
- Use an open card sort to learn how users group content and the **terms or labels they give each category**

Closed Sort

- Participants are asked to sort topics from content within your website into **pre-defined categories**
- A closed card sort works best when you are working an already fixed navigation/menu, and you want to **learn how users sort content items into each category**

Open Sort vs. Closed Sort

Example:

The content includes: “about us”, “forum”, “rating system”, “map of restaurants”, etc.

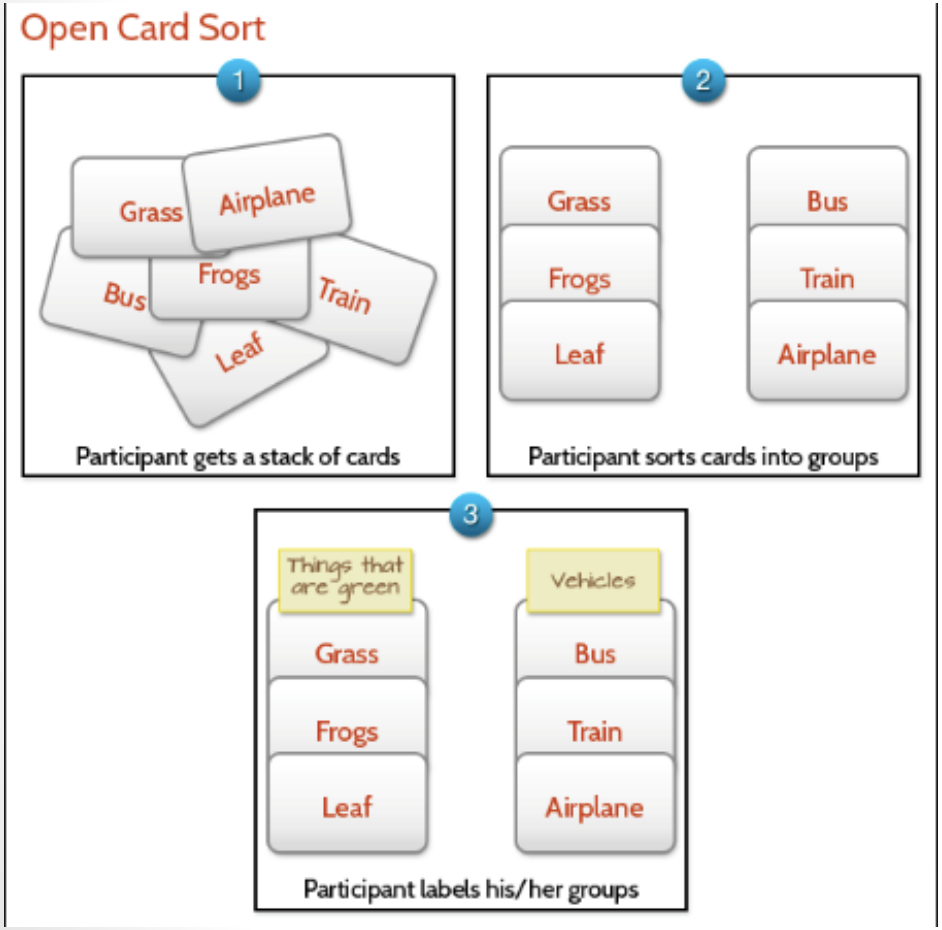
Open Sort

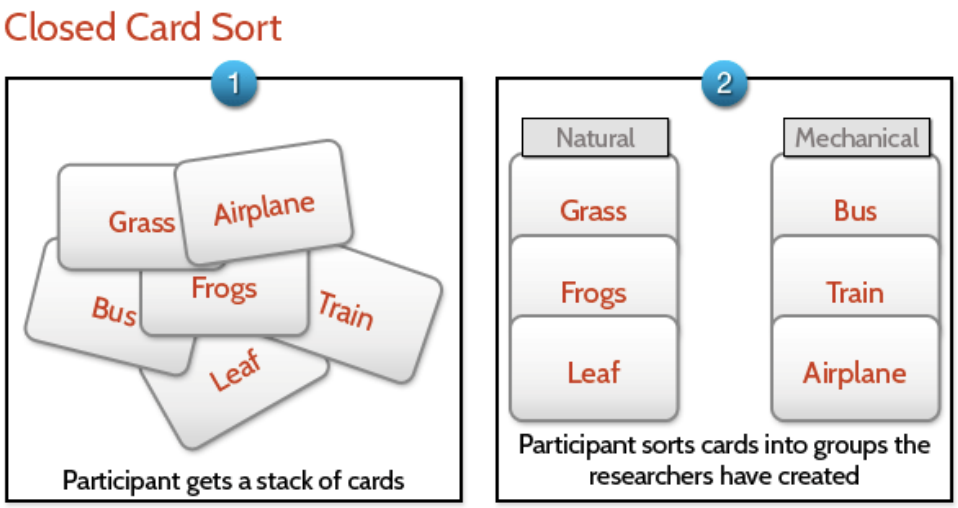
- Please organize this content into groups (no limit on # of groups)

Closed Sort

- The menu options are: “home”, “research”, “locations”, “contact us”, etc.
- Please organize content into these predefined groups

Example of Open Sort vs. Closed Sort



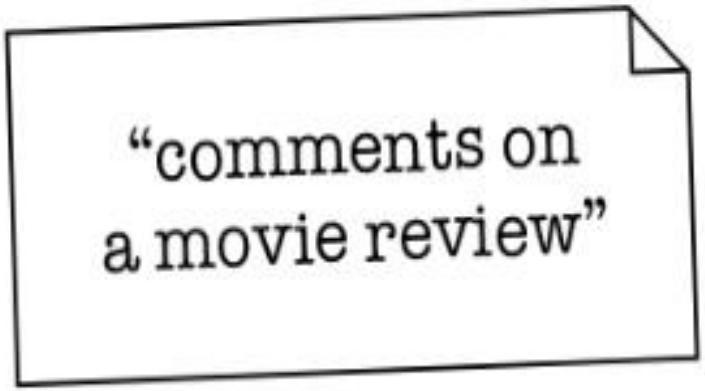


Preparing for Card Sorting

- Select content
  - Current content areas
  - Planned/future areas
  - “Blank cards” for users to create content (optional)
- Select your participants
  - Who are your users?
- Prepare the cards

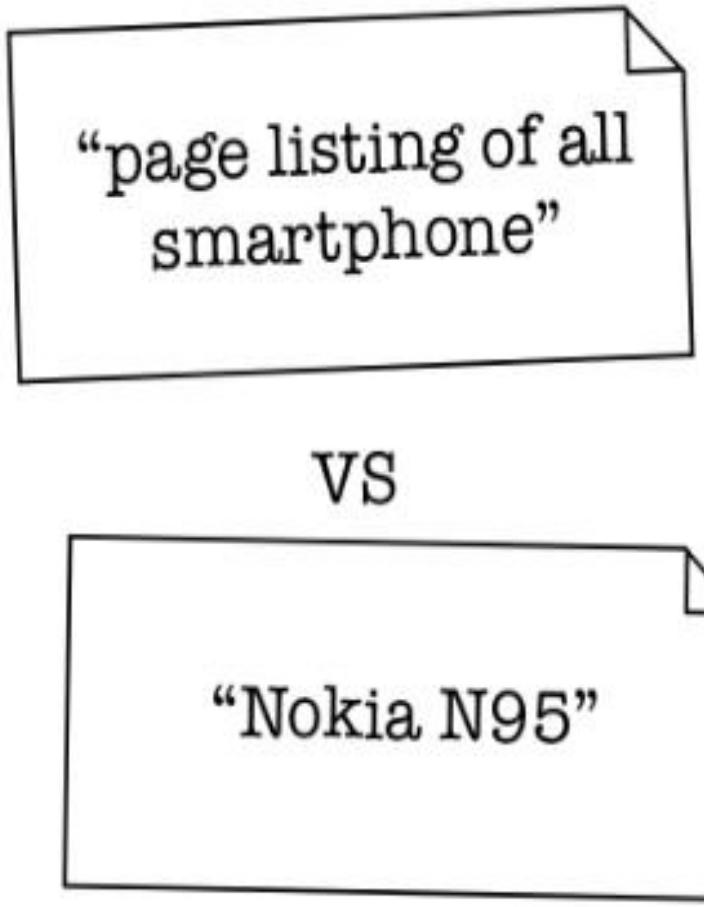
What Goes on a Card?

VS



ONE IS A PAGE, THE OTHER IS A PAGE ELEMENT

What Goes on a Card?



ONE OF THESE HAS ALREADY ENFORCED A CATEGORY...

Card Sorting Tips

1. Don’t expect the same results – discrepancies are good
2. Look for more information in the conversations than in the results (can ask participants to “think aloud”)
3. Be clear on your intentions – are validating (closed) or discovering (open)
4. Don’t equate your final card sort as your site structure --- look at the data as “input” because translation is still required
5. Run with actual users, but you can also use internally within your design team!
6. Can be conducted individually or in groups of people

More Card Sorting Tips

Detailed instructions of card sorting by Usability.gov

- https://www.usability.gov/how-to-and-tools/methods/card-sorting.html#:~:text=Card%20sorting%20is%20a%20method,help%20you%20label%20these%20groups.

Video Example

- https://youtu.be/TNvdgXCqEvM

Online Card Sorting Tools

Many online card sorting tools are available

- Free trial
- Enable both open and close card sort
- Report and analysis available

For example,

- OptimalSort
  - https://www.optimalworkshop.com/optimalsort
- UXtweak
  - https://www.uxtweak.com/card-sort-tool

Team Activity

- Continue working on prototype/testing
- Come up with your plan for usability testing
  - You've learned about several evaluation methods thus far, what evaluation methods might you use?

Reading Assignment

- ID Chapters 11, 14, and 15
- UYU Chapters 12 and 13

Evaluation Methods

Pre- & Post-prototype

- ✓Surveys: questionnaires
- ✓Surveys: interviews
- ✓Surveys: focus groups
- ✓Functional analysis
- ✓Task analysis

Post-prototype

- ✓Personas
- ✓Cognitive walkthrough
- ✓Card Sorting
- **Heuristic evaluation**
- **Field/ ethnographic**
- User testing
- Experiments

Heuristic Evaluation

Heuristic Evaluation

- Developed by Jakob Nielson

  www.nngroup.com/articles/

  - Many great resources
  - Example studies / methods
  - Literature (good place to find citations for your final projects)

Heuristic Evaluation

- Heuristic
  - Rules of thumb
  - Not guaranteed to be optimal
- Heuristic evaluation
  - Several “evaluators” <u>independently</u> critique a system using shared set of heuristics (principles or rules of thumb)

Heuristic Evaluation: The Process

- Inspect
  - Flow from screen to screen
- Evaluate against heuristics
- Find “problems”
  - Subjective & liberal (if you think it is a problem, then it is)

<u>Perform two or more passes through system</u>

Heuristic Evaluation: The Process

1. Gather inputs
2. Evaluate system
3. Severity rating
4. Debriefing

Heuristic Evaluation: Gather Inputs

- Who are the evaluators?
  - Need to learn about domain, its practices
- Prepare prototype to be studied
  - May vary from mock-ups and storyboards to a working system

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

1. Visibility of system status

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time

Does NOT mean if users can see clearly

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

2. Match between system and the real world

The system should speak the user's language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in natural and logical order

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

3. User control and freedom

Users often choose system functions by mistake and will need a clearly marked 'emergency exit' to leave the unwanted state without having to go through an extended dialog. Support undo and redo.

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

4. Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

5. Error prevention

Even better than good error messages is a careful design which prevents a problem from occurring in the first place.

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

6. Recognition rather than recall

Minimize the need of memorization

Make objects, actions and options visible. The user should not have to remember information from one part of the dialog to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

7. Flexibility and efficiency of use

Accelerators – unseen by the novice users – may often speed up the interaction for the expert users to such an extent that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

8. Aesthetic and minimalist design

Concise and relevant information

Dialogs should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialog competes with the relevant units of information and diminishes their relative visibility.

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

9. Help users recognize, diagnose, and recover from errors

Helpful error message

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

Heuristic Evaluation: Jakob Nielsen's 10 Heuristics

10. Help and documentation

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Heuristic Evaluation: Severity Rating

- Severity ratings – assign to cases where a heuristic is violated
- What do you think rating could be based on?
  - Frequency of the problem
    - Is the problem rare or common?
  - Impact of the problem
    - How hard is it for the user to overcome the deficiency?
  - Persistence of the problem
    - Can the user overcome it once and move on, or does it need to be surmounted multiple times?

Heuristic Evaluation: Severity Rating

- Severity rating is very important
  - It's how resources are allocated to the problems
  - Determines go/no-go decisions
- Inter-rater reliability in assignment of severity rankings is quite low
  - Use group consensus methods to assign severity

Heuristic Evaluation: Debriefing

- Organize all problems found by different reviewers
  - At this point, decide what are and are not major problems
  - Document and record them
  - Compare severity ratings, and determine plan of action

- Typically presented to company in written report

Example of Heuristic Evaluation

Issue/problem headers.	Heuristic violated	Severity of problem	Recommendation
Lacking column headers.	#1	2	Explicitly describe what the topic for each column is for.
BuzzFeed "reaction icons" can be foreign or unfamiliar to a new users.	#2	3	Description should be given when hovering over icon.
Promoted articles and "Preview" for posts open up in a new tab.	#3	3	While creating a post, the "Preview" should have a popup on current webpage to avoid confusion.
Inconsistency of layouts throughout website.	#4	1	Be consistent and follow the same layout for each webpage.
There is not enough white space on each webpage, it can be overwhelming.	#6	3	Reconsider certain columns to avoid repetition. Also avoid endless scrolling, offer "next" and "back" options.
Search box is not visible, the small icon can be overlooked.	#7	2	Have a larger space for the Search box area.
Overwhelming pressure to share on social media.	#8	3	Remove some social media icons, keep it less repetitive.
Two different comment sections.	#8	3	Have site comments above social media conversations, clearly describe and separate the sections.
Help and contact links not visible on the main page (under "more" section).	#10	2	The endless scrolling should be removed. Add a footer or include important help/contact links at the very top of the page. Make sure it is in a visible and appropriate section of the site.

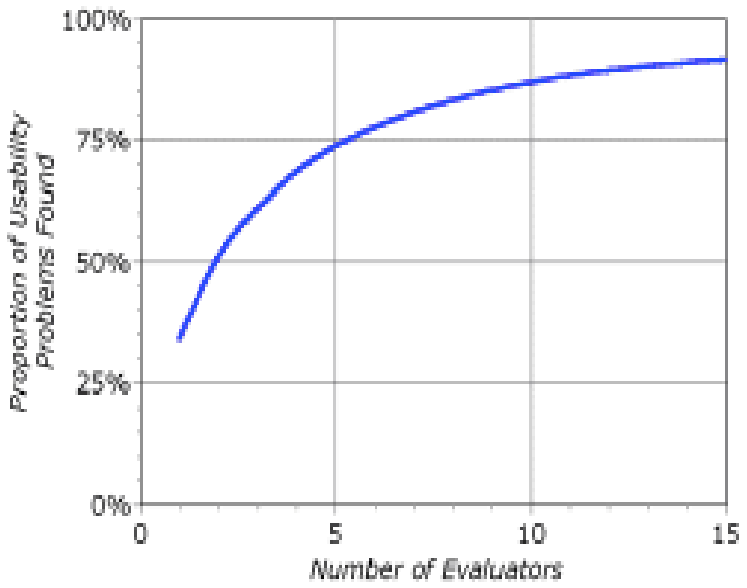
http://www.samgonzalezux.com/usability-evaluation-buzzfeed/

Heuristic Evaluation: Benefits

- One of the most common usability evaluation methods
- Highly cost effective and very fast to employ
- Easy to learn and use

Heuristic Evaluation: Weaknesses

- Need to employ more than one evaluator to get robust results.
- Want at least 3 evaluators – preferably 5
  - Results in about 75% of overall usability problems being discovered.
  - Above that get more data but less efficiency



Field Studies

Related terms: Observational Studies, Case Studies

Field / Naturalistic Observation

Observations

- Systematic assessment of overt behavior
- Natural environment

Field / Naturalistic Observation

Planning

- What, where, and when to observe / record?

  Video recordings or screen capture (if ethical)
- Taxonomy of behaviors
- Performance measures, such as
  - # of clicks
  - Time to complete task

Hawthorne Effect (observer effect)- changes in behavior that occur when people know that others are observing them

Assembly line workers

Pros / Cons of Field Studies

What do you think?

Pros and cons:

+ Large amounts of rich data
+ Capture events not duplicated in lab
+ In depth understanding
– People act differently
– Observer expectancies

User Testing

User Testing

- Often less “experimental” than normal lab studies
- Used in industry

User Testing - Procedure

- Select a set of tasks (10-30 common tasks)
- Give the user the information required to do the task
- Watch
- Record behaviors of interest

User Testing – In a Lab

Lab environment allows you to capture

- Voice
- Video
- Screen shots
- Keystrokes
- Facial expressions/body language

User Testing Metrics

As your prototype becomes more refined, usability testing becomes more quantitative.

- To collect data, a functional prototype can be built
- Users are given a set of scenarios of tasks that they would perform under usual circumstances

User Testing Metrics

The most common metrics

- Errors
- Time to perform tasks
- Time to perform subtasks
- User subjective reactions (e.g., satisfaction, preferences)

Examples of User Testing Metrics (Detail)

- **Efficiency**
  - Time to learn
  - Time spent on errors
  - Time to complete task
  - Percent or number of errors
  - Frequency of help or documentation use
  - Number of repetition of failed commands