Heuristic Evaluation

刘哲明
Prof. James A. Landay
Computer Science Department
Stanford University

Autumn 2023
November 1, 2023
Hall of Fame or Shame?

Big basket
From Bharti Bhagtani
https://uxdesign.cc/heuristic-evaluation-of-bigbasket-application-4a69f43be47d

“India’s largest online supermarket”
Big basket
From Bharti Bhagtni
https://uxdesign.cc/heuristic-evaluation-of-bigbasket-application-4a69f43be47d

“India’s largest online supermarket”

Good
- ?

Bad
- not aesthetic & minimalist design
- popups with too much info
- cluttered
Heuristic Evaluation

刘哲明
Prof. James A. Landay
Computer Science Department
Stanford University

Autumn 2023
November 1, 2023
Outline

- Heuristic Evaluation Overview
- The Heuristics
- Team Break
- Exercise (graded)
Evaluation

• About figuring out how to improve design
• Issues with lo-fi tests?
Evaluation

- About figuring out how to improve design
- Issues with lo-fi tests?

Not realistic
  - visuals & performance

Not on actual interface
  - participant can’t test alone

Need participants
  - can be hard to find repeatedly
Heuristic Evaluation

• Developed by Jakob Nielsen (now famous consulting co.)
• Helps find usability problems in a UI design
• Key idea: *small set (3-5) of evaluators* examine UI
  – independently check for compliance with usability principles ("heuristics")
  – evaluators only communicate afterwards
    • findings are then aggregated
    – use violations to redesign/fix problems
• Can perform on working UI or on sketches
Why Multiple Evaluators?

- Every evaluator doesn’t find every problem
- Good evaluators find both easy & hard to find problems
Heuristics

H1: Visibility of system status

H2: Match between system & real world

H3: User control & freedom
Heuristics (cont.)

H4: Consistency & standards

H5: Error prevention

H6: Recognition rather than recall
Heuristics (cont.)

H7: Flexibility and efficiency of use
Heuristics (cont.)

H8: Aesthetic & minimalist design

bad

good

https://icons8.com/articles/redesigning-boarding-pass-again/
H9: Help users recognize, diagnose, & recover from errors
Heuristics (cont.)

good

The page at https://mail.google.com says:
It seems like you forgot to attach a file.
You wrote *See the attached* in your message, but there are no files attached. Send anyway?

Cancel  OK
Good Error Messages

- Clearly indicate what has gone wrong
- Human readable
- Polite
- Describe the problem
- Explain how to fix it
- Highly noticeable
H10 – Help & Documentation

• Better if the system can be used without documentation, but it may be necessary

• How
  – easy to search
  – focused on task
  – list concrete steps

http://blog.screensteps.com/10-examples-of-great-end-user-documentation
H11* – Accessible Design

Users can interact with the system using *alternative input methods*. Content is legible with *distinguishable contrast* & *text size*. Key information is upfront & *not nested* for screen readers. Purely visual or auditory content has *text-based alternatives* for users with low vision & low hearing.

* New heuristic that CS147 staff has added to Nielson’s 10.
H11* – Accessible Design

Users can interact with the system using *alternative input methods*. Content is legible with *distinguishable contrast* & *text size*. Key information is upfront & *not nested* for screen readers. Purely visual or auditory content has *text-based alternatives* for users with low vision & low hearing.

* New heuristic that CS147 staff has added to Nielson’s 10.
The design should encode *values* that users can *understand and relate to*. It should make a diverse group of users feel *included and respected*. The design should *prevent the reproduction of preexisting inequities* and *not create additional burdens for disadvantaged populations*.

* New heuristic that CS147 staff has added to Nielson’s 10.
H12* – Value Alignment & Inclusion

The design should encode *values* that users can *understand and relate to*. It should make a diverse group of users feel *included and respected*. The design should *prevent the reproduction of preexisting inequities and not create additional burdens for disadvantaged populations*.

* New heuristic that CS147 staff has added to Nielson’s 10.


Welcome to Programming Methods. This course is an introduction to computer programming and software development.

ASSIGNMENTS

PROJECTS

All projects are due at 11:59 pm on the specified due date.

<table>
<thead>
<tr>
<th>Name</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard the Robots (CRIIPEND)</td>
<td>4/29</td>
</tr>
<tr>
<td>Certification Paper (PDF)</td>
<td>4/30</td>
</tr>
<tr>
<td>Mandarin (PDF)</td>
<td>5/1</td>
</tr>
<tr>
<td>Spanish (PDF)</td>
<td>5/2</td>
</tr>
<tr>
<td>English</td>
<td>5/3</td>
</tr>
<tr>
<td>Melody Maker + Image Algorithms (PDF)</td>
<td>5/5</td>
</tr>
<tr>
<td>Others (PDF)</td>
<td>5/6</td>
</tr>
<tr>
<td>JavaScript (PDF)</td>
<td>6/3</td>
</tr>
</tbody>
</table>

SECTIONS

Our sections are 50-minute discussions led by section leaders each week in which we discuss the module and work with students to solve problems. After the first week, if you are not yet in one of our CS 106a sections, you can sign up through our web site.

<table>
<thead>
<tr>
<th>Name</th>
<th>Week</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 1: Hard the Robots; Basic Java</td>
<td>4/10 - 6/10</td>
</tr>
<tr>
<td>Section 2: Cloud Computing, Network Programming</td>
<td>4/11 - 6/11</td>
</tr>
<tr>
<td>Section 3: Engineering File Processing</td>
<td>4/16 - 6/16</td>
</tr>
<tr>
<td>Section 4: Graphics, Animation, and Games</td>
<td>4/17 - 6/17</td>
</tr>
<tr>
<td>Section 5: Classes and Objects; Interactive Clones</td>
<td>4/21 - 6/21</td>
</tr>
<tr>
<td>Section 6: Arrays and Hashing</td>
<td>5/26 - 6/26</td>
</tr>
<tr>
<td>Section 7: Mathematical Casiotone Scores</td>
<td>6/20 - 6/30</td>
</tr>
<tr>
<td>Section 8: JavaScript User Interfaces</td>
<td>6/21 - 6/31</td>
</tr>
</tbody>
</table>

Are you looking for the list of section leader names and section times? That information can be found on the web site, but only for the current quarter.

ASSIGNMENTS

PROJECTS

All projects are due at 11:59 pm on the specified due date.

<table>
<thead>
<tr>
<th>Name</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard the Robots (CRIIPEND)</td>
<td>4/29</td>
</tr>
<tr>
<td>Certification Paper (PDF)</td>
<td>4/30</td>
</tr>
<tr>
<td>Mandarin (PDF)</td>
<td>5/1</td>
</tr>
<tr>
<td>Spanish (PDF)</td>
<td>5/2</td>
</tr>
<tr>
<td>English</td>
<td>5/3</td>
</tr>
<tr>
<td>English</td>
<td>5/4</td>
</tr>
<tr>
<td>Spanish</td>
<td>5/5</td>
</tr>
<tr>
<td>Mandarin</td>
<td>5/6</td>
</tr>
<tr>
<td>Spanish</td>
<td>5/7</td>
</tr>
<tr>
<td>Mandarin</td>
<td>5/8</td>
</tr>
<tr>
<td>Others (PDF)</td>
<td>5/9</td>
</tr>
<tr>
<td>JavaScript (PDF)</td>
<td>6/3</td>
</tr>
</tbody>
</table>

SECTIONS

Our sections are 50-minute discussions led by section leaders each week in which we discuss the module and work with students to solve problems. After the first week, if you are not yet in one of our CS 106a sections, you can sign up through our web site.

<table>
<thead>
<tr>
<th>Name</th>
<th>Week</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 1: Hard the Robots; Basic Java</td>
<td>4/10 - 6/10</td>
</tr>
<tr>
<td>Section 2: Cloud Computing, Network Programming</td>
<td>4/11 - 6/11</td>
</tr>
<tr>
<td>Section 3: Engineering File Processing</td>
<td>4/16 - 6/16</td>
</tr>
<tr>
<td>Section 4: Graphics, Animation, and Games</td>
<td>4/17 - 6/17</td>
</tr>
<tr>
<td>Section 5: Classes and Objects; Interactive Clones</td>
<td>4/21 - 6/21</td>
</tr>
<tr>
<td>Section 6: Arrays and Hashing</td>
<td>5/26 - 6/26</td>
</tr>
<tr>
<td>Section 7: Mathematical Casiotone Scores</td>
<td>6/20 - 6/30</td>
</tr>
<tr>
<td>Section 8: JavaScript User Interfaces</td>
<td>6/21 - 6/31</td>
</tr>
</tbody>
</table>
Experimental Measures

<table>
<thead>
<tr>
<th>Enrollment intentions</th>
<th>Sense of belonging</th>
<th>Anticipated Success</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technical self-confidence</th>
<th>Future CS intentions</th>
<th>Stereotype anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Neutral
Stereotypical

Heuristic Violation Examples

1. [H6 Recognition Rather Than Recall]
   Can’t copy info from one window to another
   - user needs to memorize the data & retype
   - fix: allow copying

2. [H4 Consistency and Standards]
   Typography uses different fonts in 3 dialog boxes
   - slows users down
   - probably wouldn’t be found by user testing
   - fix: pick a single format for entire interface
Severity Ratings

0 - don’t agree that this is a usability problem
1 - cosmetic problem
2 - minor usability problem
3 - major usability problem; important to fix
4 - usability catastrophe; imperative to fix
1. [H4 Consistency & Standards] [Severity 3]

The interface used the string “Save” on the first screen for saving the user’s settings, but used the string “Store” on the second screen. Users may be confused by this different terminology for the same function.

Fix: Use “Save” everywhere in the application.
Cart

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Item #</th>
<th>Remove?</th>
<th>Item Description</th>
<th>Price</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>#033DVM4800</td>
<td>□</td>
<td>Denon DVM-4800 DVD Video Changer</td>
<td>$1,199.95</td>
<td>$1,199.95</td>
</tr>
<tr>
<td>1</td>
<td>#204EL570</td>
<td>□</td>
<td>Bazooka EL570 5&quot;x7&quot; 2-way speakers</td>
<td>$79.95</td>
<td>$79.95</td>
</tr>
<tr>
<td>1</td>
<td>#158DSCP50</td>
<td>□</td>
<td>Special Sony DSC-P50 Digital Still Camera, 2.1 Mega Pixel</td>
<td>$299.95</td>
<td>$299.95</td>
</tr>
<tr>
<td>1</td>
<td>#123DVDLA95</td>
<td>□</td>
<td>Panasonic DVD-LA95 Portable DVD-AV Player with 9&quot; Screen</td>
<td>$999.95</td>
<td>$999.95</td>
</tr>
<tr>
<td>1</td>
<td>#170GXT160</td>
<td>□</td>
<td>Maxell GX-Silver T-160 VHS Video Tape</td>
<td>$2.49</td>
<td>$2.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Merchandise Total: **$2,582.29**

Standard Shipping Charge: **$13.95**
Alternative shipping options available before final checkout
Order Total: **$2,596.24**

- To change an item's quantity, enter the correct number in the Quantity column, then press Update Cart.
- To remove an item, check the box in the Remove? column, then press Update Cart.
- To order an item that appears in your printed Crutchfield catalog, enter the item number into the Cart and click Update Cart.
- International visitors, please click here.
Problems Found this Year

- Foo
Problems Found this Year

- Bar
Problems Found this Year

- Baz
Problems Found Last Year

- **H1: Visibility of system status**
  - Username for current shopping cart missing [1]
- **H2: Match between system & real world**
  - "price" and "total" should refer to different things [1]
  - Items referred to by their Item Number [3]
- **H3: User control & freedom**
  - Not clear where/how user can place order [1]
  - Missing a search bar [1]
- **H4: Consistency & standards**
  - Red used inconsistently - error, special, title, out of stock [4]
  - Yes/No Checkbox in “Remove?” [5]
  - Menu Bar links on different lines - logout, continue, shopping [1]
Problems Found Last Year

- **H5: Error prevention**
  - Out of Stock items allowed to be added to cart [4]
  - ”h” in Quantity Field [7]
- **H6: Recognition rather than recall**
  - Item number not auto-filled, difficult for user to remember [2]
- **H7: Flexibility and efficiency of use**
  - Update Cart Button required to update cart and ambiguous use [2]
  - Quantity Field could be a selection menu [1]
- **H8: Aesthetic & minimalist design**
  - Color Combinations - red/green, black text on blue background [6]
  - Crowded Design - columns and text squished [2]
  - # Sign not necessary [1]
Problems Found Last Year

- H9: Help users recognize, diagnose, & recover from errors
  - “Please Note” error message refers to item bolded in red but multiple items bolded in red [2]
  - Last row has quantity 1 but no other info, user unsure how to fix [1]
- H10: Help & Documentation
  - Instructions for use not immediately visible [1]
  - International users not immediately redirected to other page [2]
- H11*: Accessible
  - Aesthetic choices difficult for visually impaired- blue links, small text, blue on blue color scheme [3]
  - Error message relies on ability to view red color - hard for color-blind [1]
- H12*: Fairness & inclusion
  - Requires knowledge of your car beyond just model and year [1]
Problems Found Two Years Ago

- H4 (consistency): The use of red to indicate errors, out of stock items, and sales/savings [26]
- H4 (consistency): two of the check boxes have yes/no next to them but none of the others do. Error prevention? [22]
- H9 (Aesthetic): "#" is not needed [2]
Decreasing Returns

Proportion of Usability Problems Found

Number of Evaluators

Benefits / Cost

Number of Evaluators

* Caveat: graphs for a specific example
Heuristic Evaluation Summary

- Have evaluators go through the UI twice
- Ask them to see if it complies with heuristics
  - note where it doesn’t & say *why*
- Have evaluators independently rate severity
- Combine the findings from 3 to 5 evaluators
  - come to agreement on problems, fixes & severity

- Alternate with user testing – you’ll find different types of problems
Speech UI Heuristics

S1: Give the agent a persona through language, sounds, and other styles.

S2: Make the system status clear.

S3: Speak the user’s language.

S4: Start and stop conversations.

S5: Pay attention to what the user said and respect the user’s context.

S6: Use spoken language characteristics.

S7: Make conversation a back-and-forth exchange.

S8: Adapt agent style to who users are, how they speak, and how they are feeling.

S9: Guide users through a conversation so they are not easily lost.

S10: Use responses to help users discover what is possible.

Evaluating Speech-Based Smart Devices Using New Usability Heuristics

Zhiwen Wen
designing AI
Stanford University

We developed a set of 17 usability heuristics for speech-based smart devices. An expert evaluation of three popular devices shows that these heuristics can be used to uncover existing usability problems as well as help design new interfaces.

A recent empirical study showed that in both English and Mandarin, speaking is more than three times faster than typing a short message. Thanks to recent breakthroughs in speech and language technologies, speech user interfaces (SUIs) have appeared rapidly on voice-enabled devices. However, designing a good SUI system remains challenging. The state of SUI design is often opaque to users, leading to user utterances compared in graphical user interfaces (GUIs). Unfortunately, simplifying GUIs into voice interfaces does not work well. Although researchers have been working on SUI technology for three decades, most of this knowledge is still hidden in technical papers and not easily accessible to designers. Moreover, the knowledge has not been applied to reflect recent improvements in speech-recognition accuracy. Consequently, these new SUI design efforts have been slow.

To help address these issues, we developed a new set of heuristics for designing and evaluating speech-based smart devices. To validate and improve these heuristics, we tested a group of usability experts—half of whom specialized in SUIs —and other groups to empirically evaluate three state-of-the-art devices.

RELATED WORK

In the early 1990s, Jakob Nielsen developed a set of 10 usability heuristics for evaluating UIs (www.useit.com/about-us/usability-heuristics). Although these heuristics are most often applied to GUIs, he and his colleagues also used them to evaluate a telephone voice-activated system. Here, the worst aspect of the system was quite limited.

Published by the IEEE Computer Society
Speech UI Heuristics

S11: Keep feedback and prompts short.
S12: Confirm input intelligently.
S13: Use speech-recognition system confidence to drive feedback style.
S14: Use multimodal feedback when available.
S15: Avoid cascading correction errors.
S16: Use normal language in communicating errors.
S17: Allow users to exit from errors or a mistaken conversation.

The list of heuristics along with detailed descriptions and examples can be found at http://hci.stanford.edu/publications/2018/speech-he/sui-heuristics.html
Administrivia

- Individual heuristic evaluation assignment next week

- Midterm two weeks from today (in class plus at home design problem)
  - Have an OAE letter? If you haven’t gotten it to us, you must by the end of today or we will not able to accommodate
Grading on Last Assignment

Assignment #5 Low-fi Prototype & Test

A5 Group Grade: ✓ --: 3% ✓ -: 8% ✓ : 56% ✓ +: 33% ✓ ++: 0%
A5 Individual Presentation: ✓ --: 0% ✓ -: 3% ✓ : 18% ✓ +: 70% ✓ ++: 8%

A5 Group Average: 91%
A5 Individual Presentation Average: 94%
TEAM BREAK
EXERCISE

Heuristic evaluation practice
(graded – turn it in at the end of class)
[8 min]: On your own, find 8+ usability violations (type in a google doc/write on this paper & circle things on this image – write your name & SUNetID)

For each one: Which guideline was violated and why? How could it be fixed?

[5 min]: Share with your group (3-4)
What violations did you all find? Which ones did only one of you notice?

[5 min]: Share with the class
Raise your hand or post a screenshot with your description on Slack.

Find, label, & describe 8-10 Heuristic Violations
[10 min]: Share with the class
Raise your hand or post a screenshot with your description on Slack. **Turn in** with your name & SUNetID on piece of paper at end of class (by studio).

- **H8 - low contrast**
  - Choose Recipient

- **H2, H8**
  - User selects James L. and presses "NEXT"

- **H6**
  - Where is the recipient?

- **H3**
  - User types message in the yellow box, and presses "Send Text"

- **H1**
  - Can only go one direction

- **H3**
  - Where is the message confirmation?

- **H3**
  - (NULL)

- **H3**
  - Create Message >

- **H8**
  - Why are these buttons all different? H3

- **H8**
  - How can I go back?
Problems Found Last Year

**H2: Match between system & real world**
Showing “(null)” for empty content

**H3: User control & freedom**
No “back” button [5]

**H6: Recognition rather than recall**
No way to see who you’re sending to in the second screen
Problems Found Two Years Ago

1. H1: no feedback on whether the message is sent or not, just navigates to inbox screen
2. [H3: user control and freedom] No way to cancel message solution: add an ‘x’ or cancel button that allows them to delete the message on screen 2
3. H8. There was unnecessary users listed that were null
4. H4: so many different colors, not consistent
5. H1: No visibility of system status (can’t really tell which page I’m on - perhaps some sort of diagram at the bottom could fix this)
6. H6: James L is not displayed on create message screen
7. [H6 Recognition rather than recall] Can’t see who you’re sending the message to when you’re creating the message. Fix: include sub-header that specifies name you just clicked on
8. H4: inconsistent button design
9. H4: consistency with names on the first screen (names have or don’t have last initials) --> should just choose one format (potentially just first name if we want to avoid clutter)
10. H8 (aesthetics) - The headers have low contrast (especially the purple/pink) and should use less saturated colors for better contrast.
11. H3/H6: User can’t go back from screen 2 to 1 if they misclicked on the wrong person’s profile, no recall of who you selected --> recall the recipient name on screen 2
12. H3: User control and freedom - can only send to recipients in the list, no way to add a new recipient
13. H9: “error”/null messages offer no explanation
14. H3: can only send text (no images/files which might be helpful in certain cases)
15. [H5 Error prevention] Could potentially select (null) users to write messages to. Fix: just don’t display the icons of null users at all.
16. H7: unnecessary “next” button on the first page. Might instead click on the person to transition to the next screen, and display the person’s name on screen 2 (H6) and add back button there (H3)
17. H3: How do you view old messages?
18. H10: no help or documentation included - just buttons indicating the very next step
19. H4 consistency: not sure the differences between ‘text’ ‘messages’ ‘body’
20. H3 - no ability to edit message text
21. H2: “Body” and “Null” might not make sense to user without technical background, but also why would they even see these words to begin with
22. H4: terminology of “message” vs “text”
Further Reading

Heuristic Evaluation

- Books
  - *Usability Engineering*, by Nielsen, 1994

- Web site
  - [http://www.nngroup.com/articles/](http://www.nngroup.com/articles/)

- Accessibility
  - [Accessibility - Foundations - Human Interface Guidelines - Design - Apple Developer](http://www.nngroup.com/articles/)

Autumn 2023
Exit Ticket

Next Time

- Designing the Future: Early and Future Visions of HCI
  - Read
    - As We May Think by Vannevar Bush
    - Tools For Thought (Ch 9), Engelbart Demo
  - Listen and Read: Of Mice and Men, 99% Invisible, Episode 149 (21 minutes)

- Studio
  - Half-way review of medium-fi with outside experts