

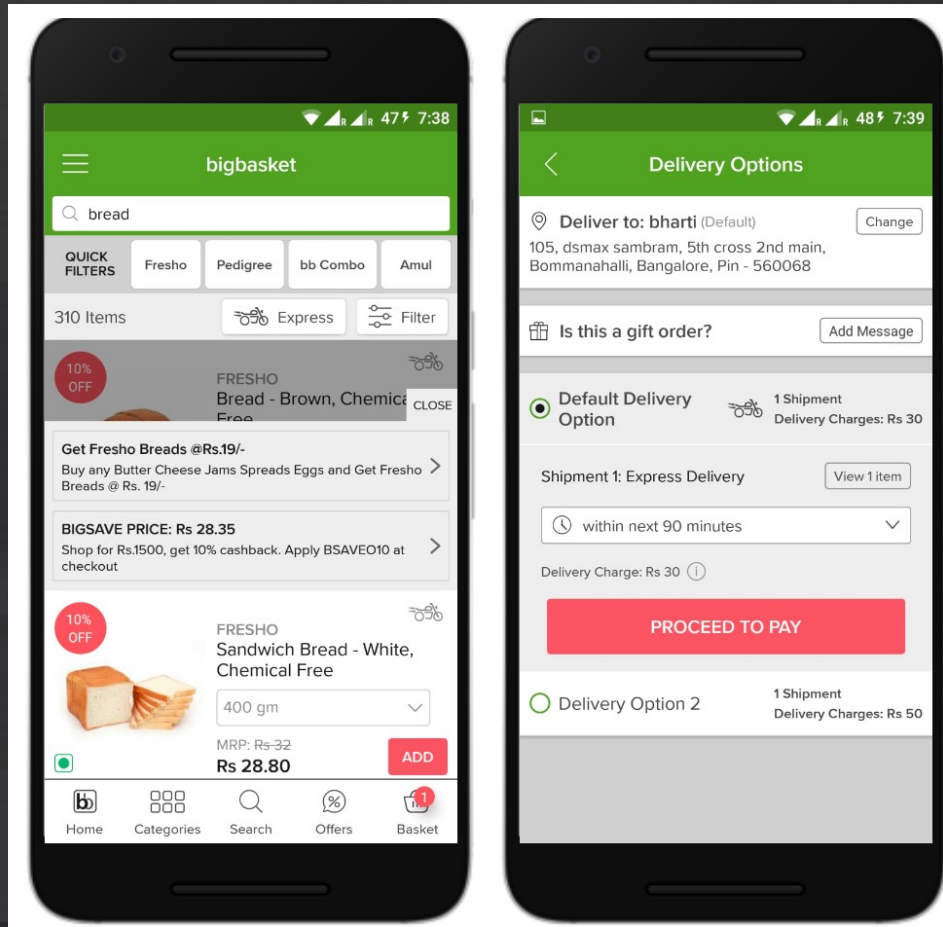
Heuristic Evaluation

刘哲明

Prof. James A. Landay
Computer Science Department
Stanford University

Autumn 2023
November 1, 2023

Hall of Fame or Shame?



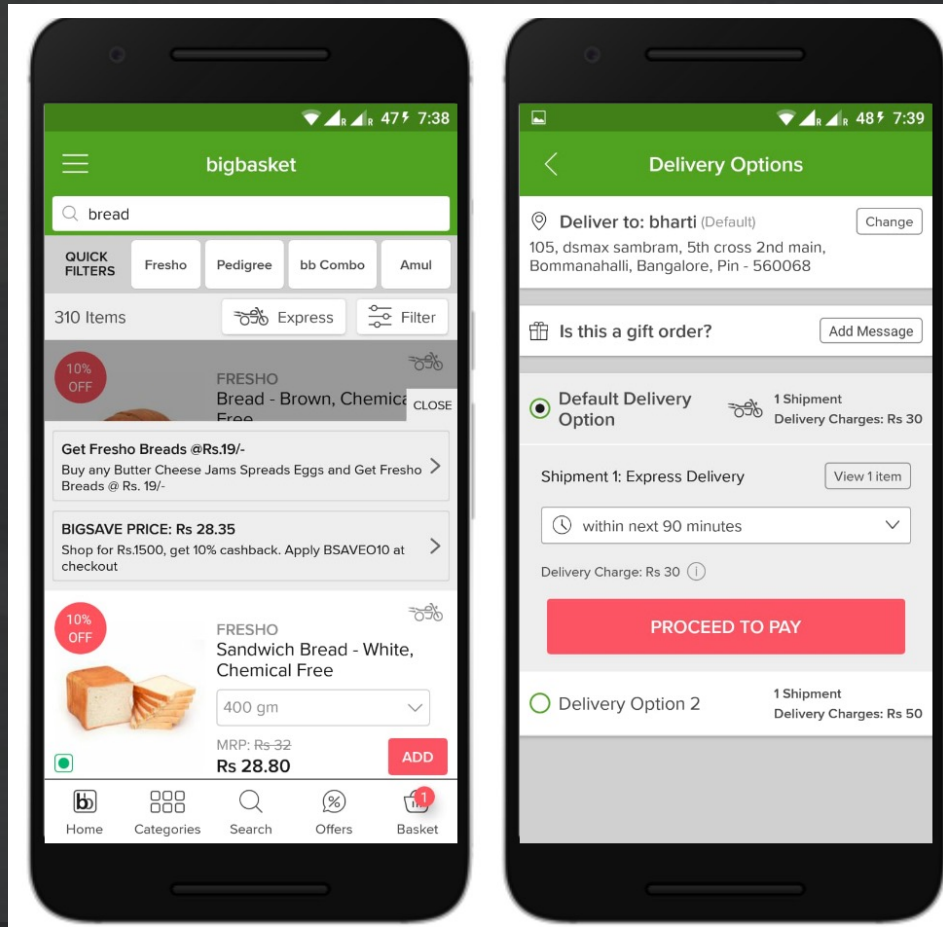
Big basket

From Bharti Bhagani

<https://uxdesign.cc/heuristic-evaluation-of-bigbasket-application-4a69f43be47d>

“India’s largest online supermarket”

Hall of Shame!



Big basket

From Bharti Bhagani

<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

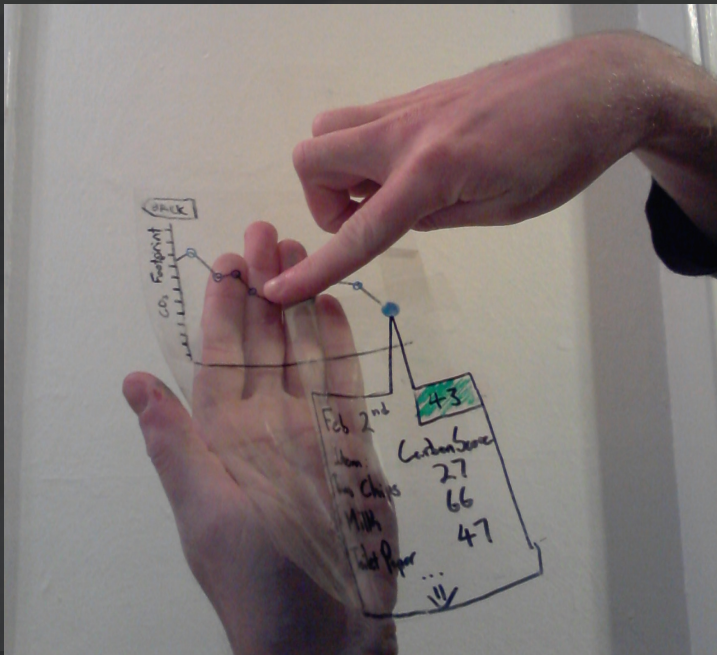
Autumn 2023
November 1, 2023

Outline

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

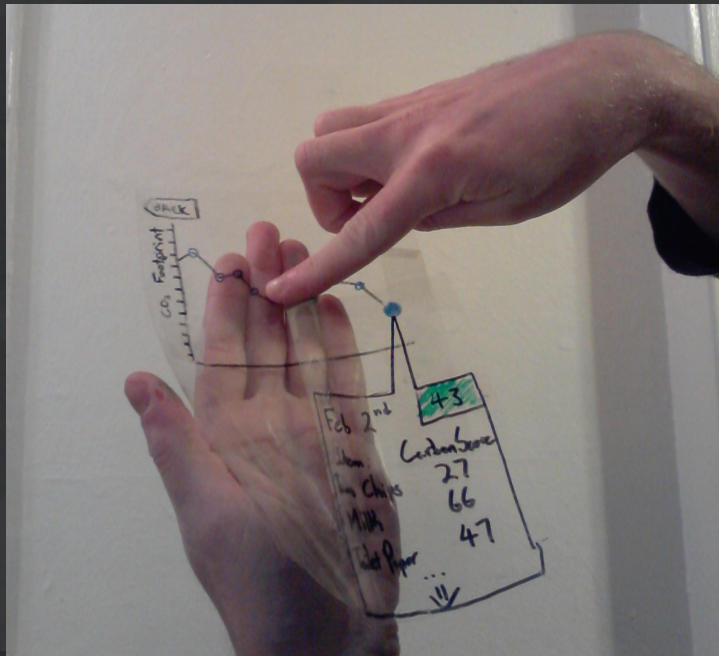
Evaluation

- 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

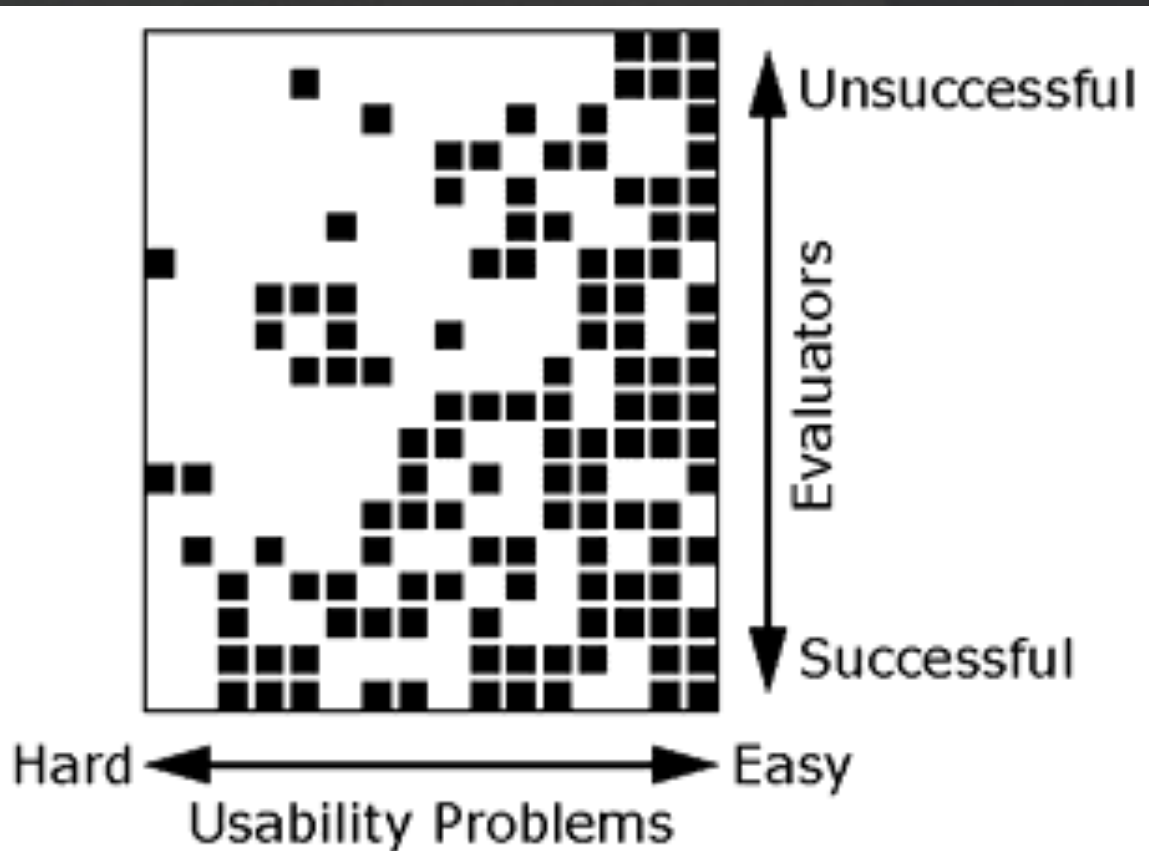
Need participants

- 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?

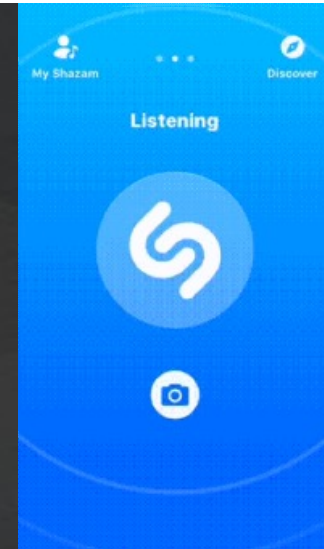
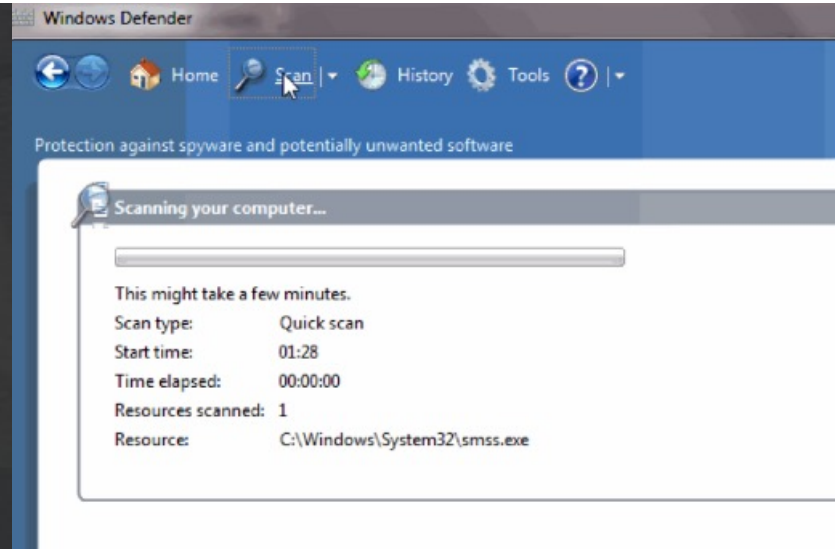


- Every evaluator doesn't find every problem
- Good evaluators find both easy & hard to find problems

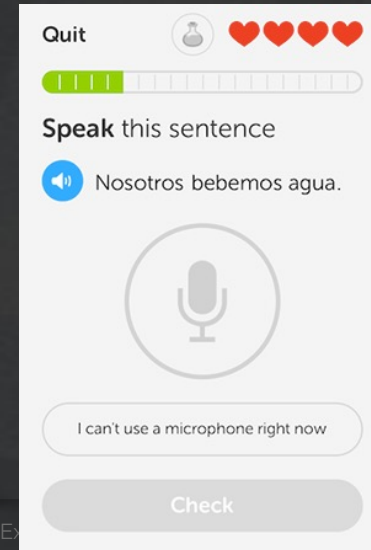
Heuristics

H1: Visibility of system status

H2: Match between system & real world

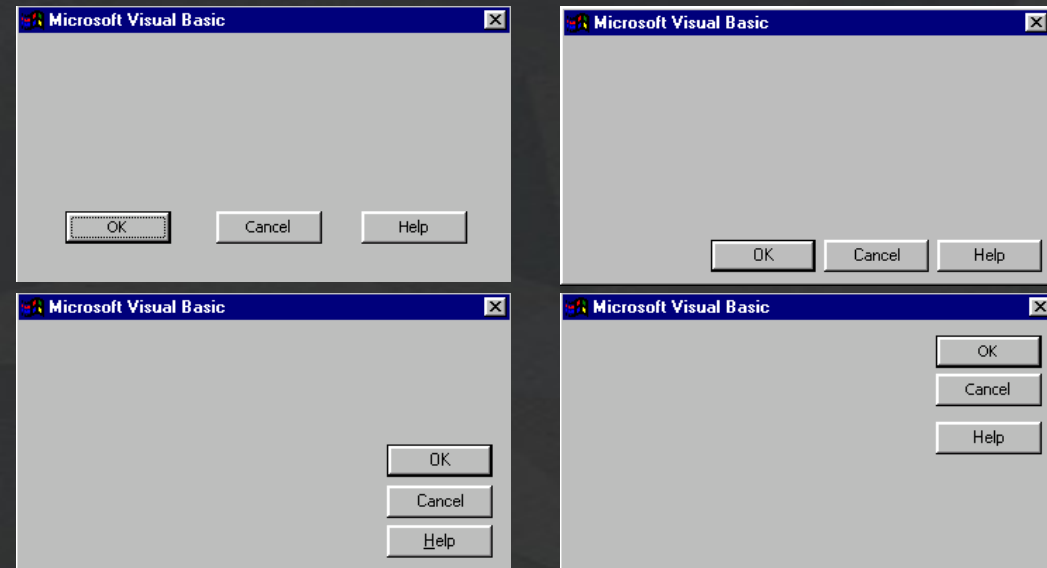


H3: User control & freedom



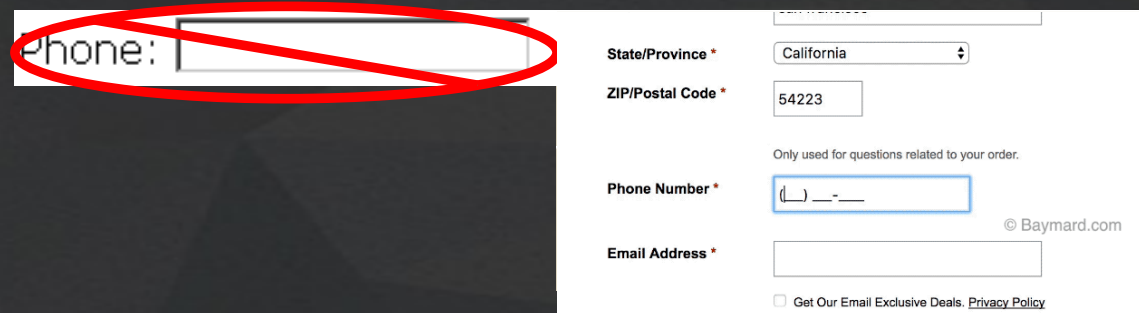
Heuristics (cont.)

H4: Consistency & standards



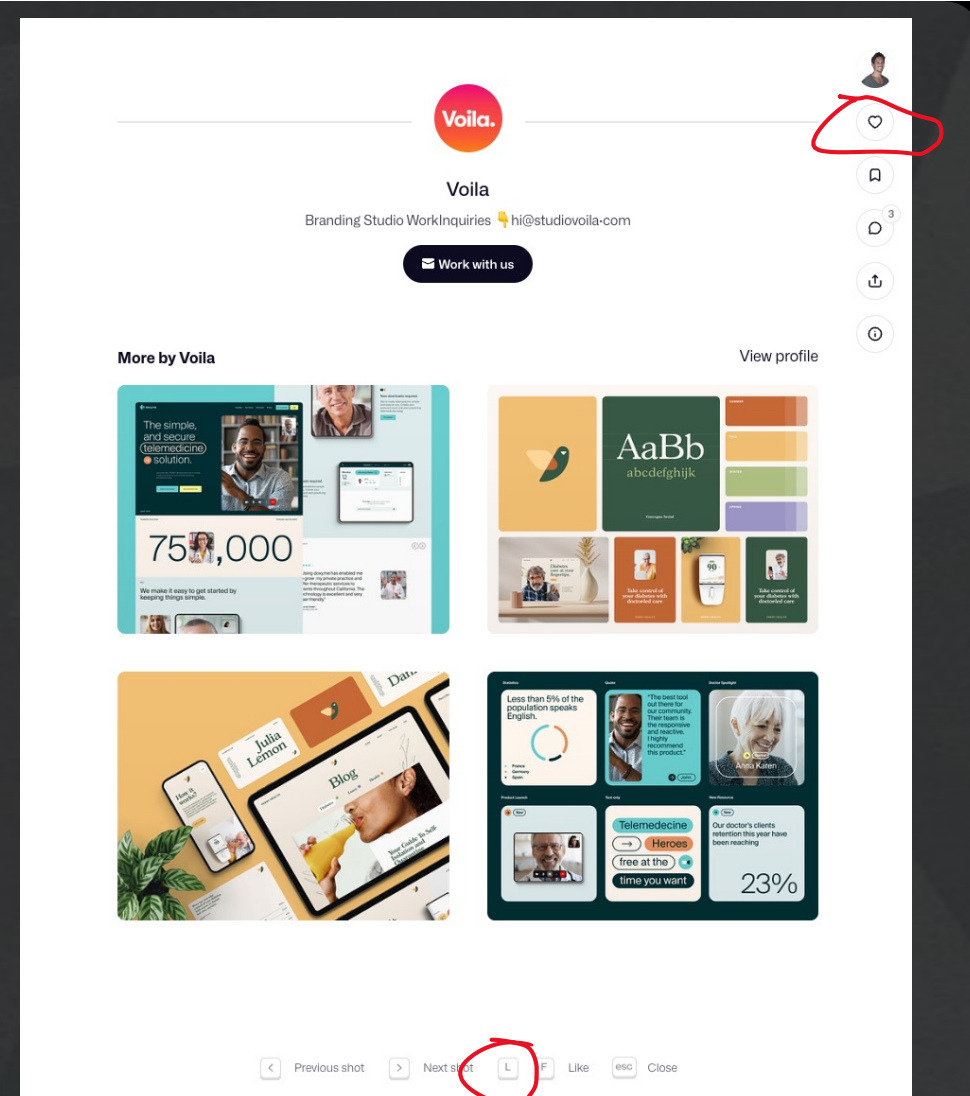
H5: Error prevention

H6: Recognition rather than recall



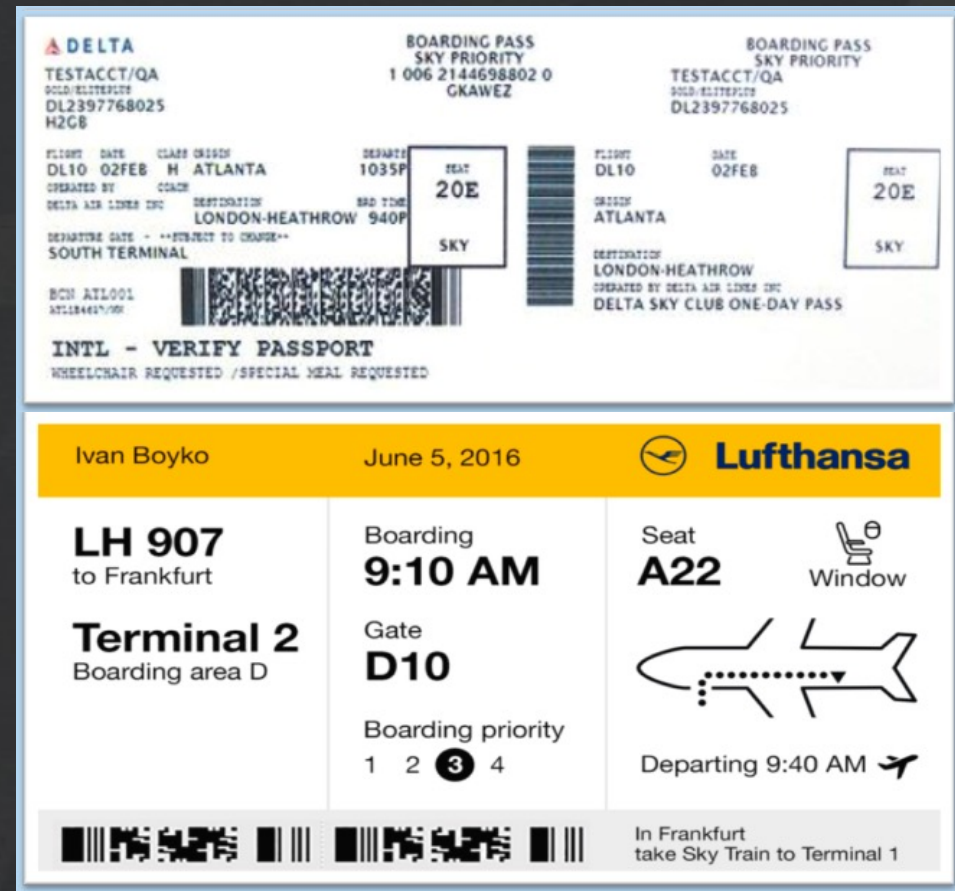
Heuristics (cont.)

H7: Flexibility and efficiency of use



Heuristics (cont.)

bad



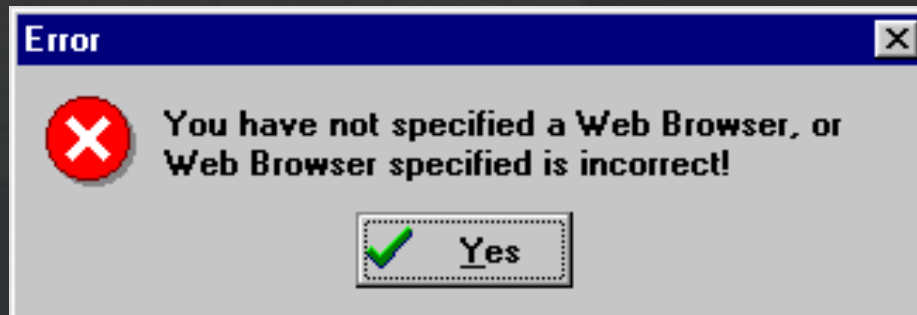
good

H8: Aesthetic & minimalist design

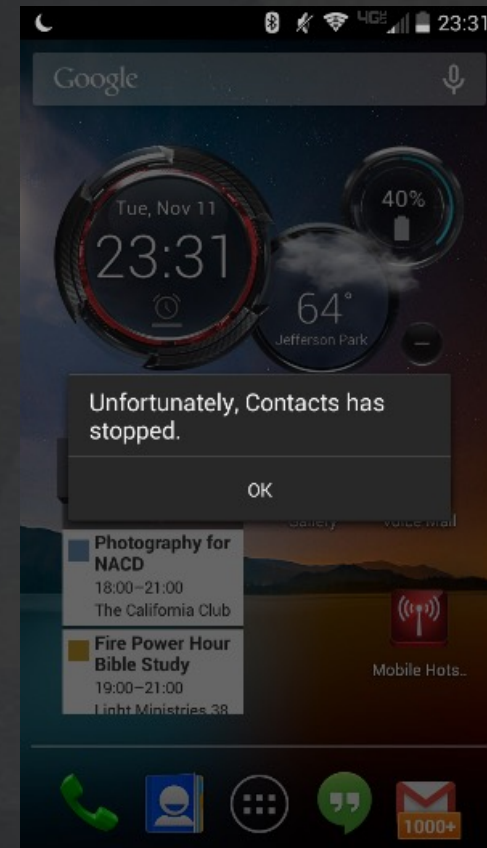
<https://icons8.com/articles/redesigning-boarding-pass-again/>

Heuristics (cont.)

bad

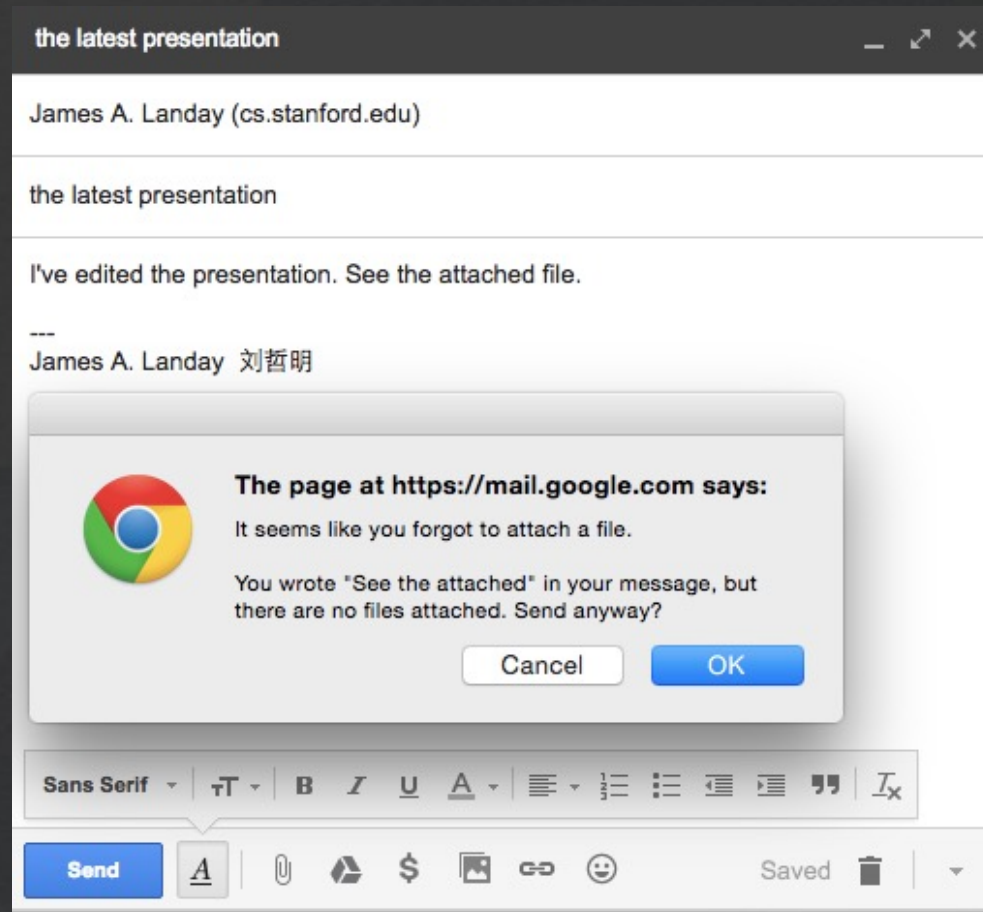


H9: Help users recognize, diagnose, & recover from errors



Heuristics (cont.)

good



Good Error Messages

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

The screenshot shows a car insurance quote page for 'Tyre & Auto Service'. The page is titled 'Your premium estimate' and displays a premium of \$443* for a 2003 Holden Commodore. A red error message box is highlighted with a green circle containing the number '1'. The error message asks the user to answer two questions: 'Does the car have any non-standard accessories or modifications?' and 'When would you want your new policy to start?'. Below the error message, there is a section for 'Get an exact quote' with a dropdown menu for 'What type of insurance are you looking for?' set to 'Comprehensive'. A second red error message box is highlighted with a green circle containing the number '2'. This error message is triggered by the date selection (28 February 2011) and asks the user to choose a start date within the next 30 days. The page also includes a 'Need help with your quote?' section with a phone number, a 'CAR INSURANCE from the car experts' header, and a sidebar with 'Information you have provided' including car details, kept overnight status, main driver information, and cover for drivers under 25.

Tyre & Auto Service

Need help with your quote?
Call us on 1300 590 766

CAR INSURANCE
from the car experts

Your premium estimate

To continue, please...

Answer these questions:

- [Does the car have any non-standard accessories or modifications?](#)

Correct the answer to these questions:

- [When would you want your new policy to start?*](#) (Please choose a start date that is within the next 30 days.)

For one year of comprehensive car insurance, your premium will be about **\$443***
* May change depending on your particular circumstances

The amount above is just an estimate and may change depending on your particular circumstances. If you provide a little more information below, we can tell you exactly what your premium will be.

Get an exact quote

Note that we need answers for all of these questions to be able to provide you with an exact quote.

What [type of insurance](#) are you looking for? Comprehensive

When would you want your new policy to start? 28 February 2011
Please choose a start date that is within the next 30 days.
Your policy can start any time in the next 30 days

Who currently insures your car? AAMI

Information you have provided

Car: 2003 Holden Commodore S Manual 4D Sedan
[Change](#)

Kept overnight: In postcode 2607
[Change](#)

Main driver: 32 year old male
made no car insurance claims in the last 3 years
[Change](#)

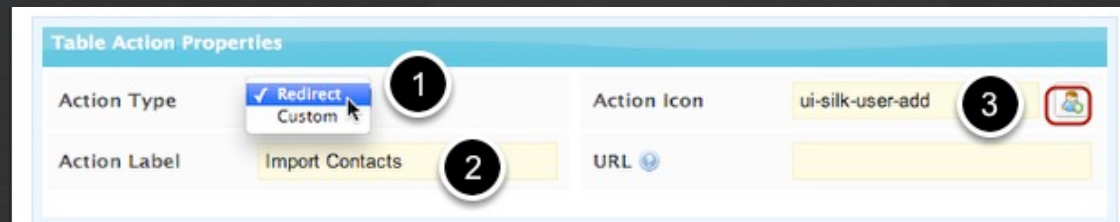
Cover for drivers under 25? No
[Change](#)

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



The screenshot shows a 'Table Action Properties' dialog box with the following fields and values:

Table Action Properties	
Action Type	<input checked="" type="checkbox"/> Redirect <input type="checkbox"/> Custom
Action Label	Import Contacts
Action Icon	ui-silk-user-add
URL	

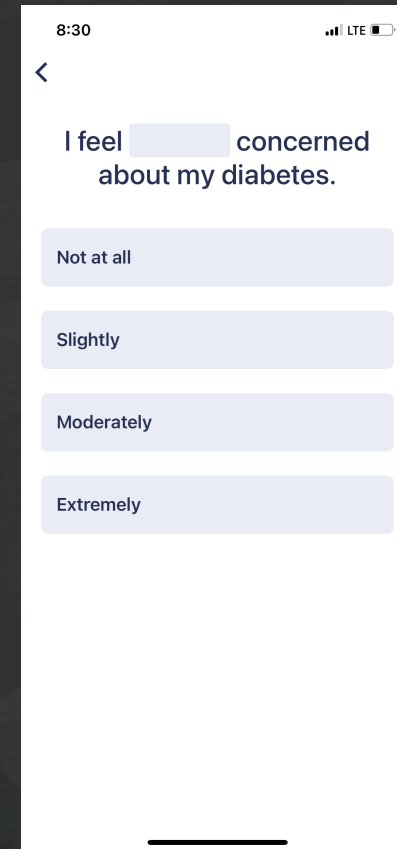
Numbered callouts: 1 points to the 'Redirect' option in the Action Type dropdown; 2 points to the 'Import Contacts' text in the Action Label field; 3 points to the 'ui-silk-user-add' icon in the Action Icon field.

1. Choose **Redirect** as the action type.
2. Name your action something inspirational, like **Import**.
3. Choose a helpful icon.

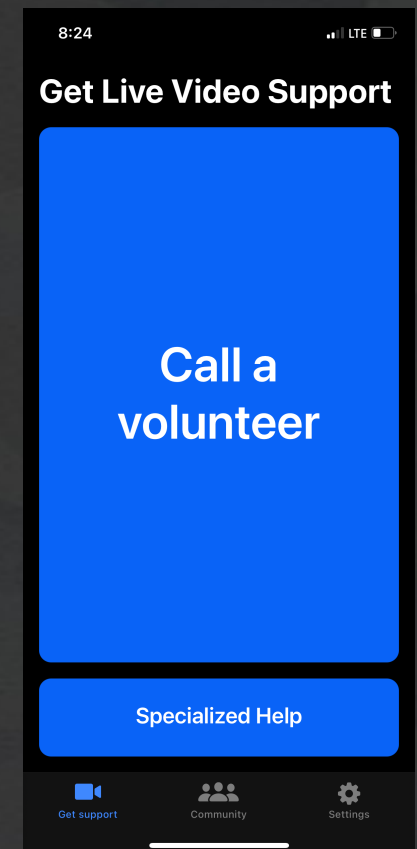
<http://blog.screensteps.com/10-examples-of-great-end-user-documentation>

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.



Glucose Buddy

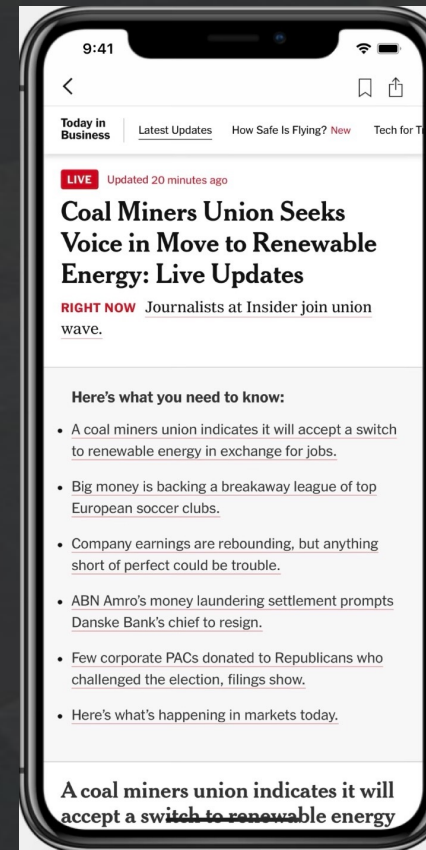


Be My Eyes

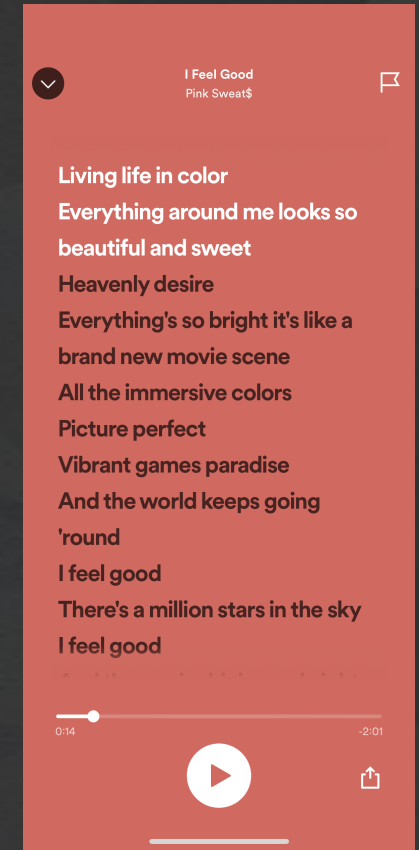
* New heuristic that CS147 staff has added to Nielsen's 10.

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.



NY Times



Spotify (lyrics mode)

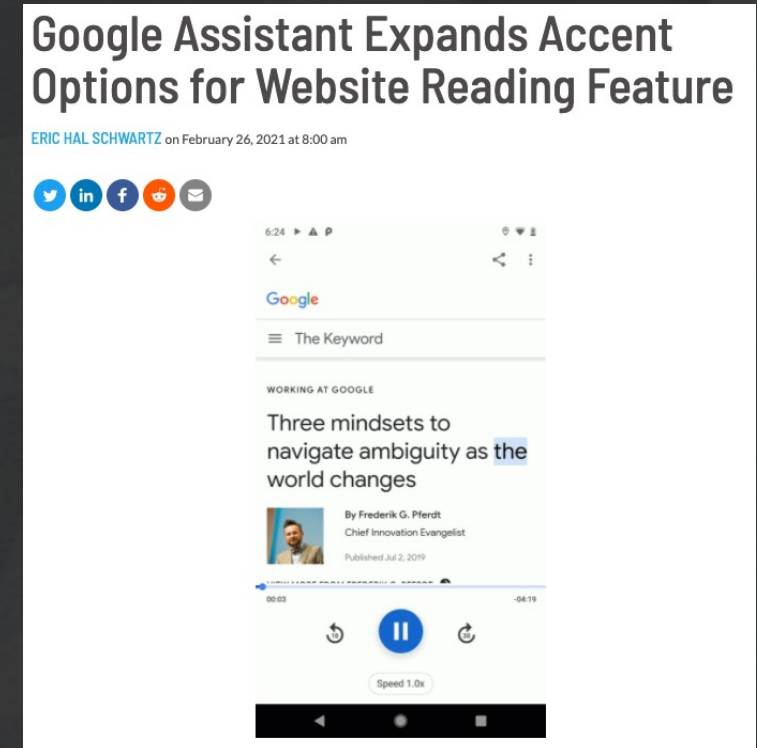
* New heuristic that CS147 staff has added to Nielson's 10.

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*.



Google adds accented reading but leaves out marginalized groups.

* New heuristic that CS147 staff has added to Nielsen's 10.

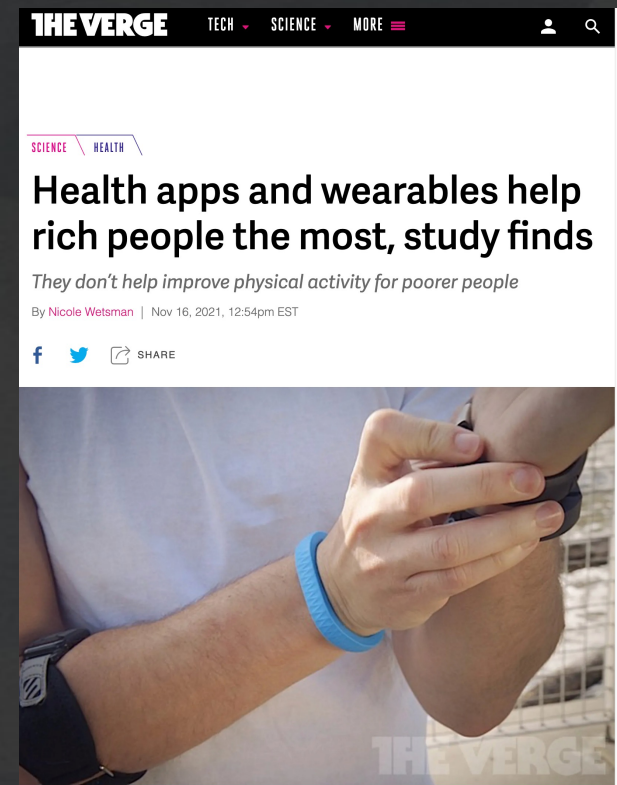
H12* – Value Alignment & Inclusion

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 Nielsen's 10.



<https://www.theverge.com/2021/11/16/22785163/digital-health-physical-activity-inequity>

Western, M.J., Armstrong, M.E.G., Islam, I. *et al.* The effectiveness of digital interventions for increasing physical activity in individuals of low socioeconomic status: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 18, 148 (2021). <https://doi.org/10.1186/s12966-021-01218-4>

HOME LECTURES ASSIGNMENTS SECTION PIAZZA TA HOURS AND STAFF

WELCOME TO CS15!
INTRODUCTION TO OBJECT-ORIENTED PROGRAMMING

HOME ASSIGNMENTS HOURS LECTURES DOCUMENTS STAFF

CS22

Introduction to Discrete Structures and Probability

CS22 GIVES YOU THE TOOLS TO SOLVE INTERESTING PROBLEMS.

Ever wanted to construct a solid, bullet-proof argument? Felt the need to count really large things? Wondered about the math behind spam filters and RSA cryptography? If so, let's get to work on Probability, Combinatorics, Logic, Graph Theory, and more!

You will learn to see the world differently, no longer accepting what is presented to you, but instead questioning, building, and exploring.

CS8: A First Byte

News Schedule Assignments Labs TA Hours Course Info FAQ Staff

CS8:
A FIRST BYTE OF
COMPUTER SCIENCE

CS166 / CS162
Computer Systems Security

Home Assignments Lectures Calendar Demos Staff Resources

Welcome to CS 166 and 162!
Register for CS166 on Piazza [here](#).
Join the waitlist [here](#).

About CS 166
CS 166 meets TuTh 1-2:20 in CIT 368. The first class is on Thursday, January 25.

CS166 teaches principles of computer security from an applied viewpoint and provides hands-on experience with security threats and countermeasures. The course additionally covers principles and skills useful for making informed security decisions and for understanding how security interacts with the world around it. The main topics covered are cryptography, authentication, access control, operating systems security, web security, and network security. Other topics include general security principles, human factors such as trust and social engineering, security of complex systems, and the economics of security. The course aims to balance theory and practice.

CS051 Models of Computation
The Next Generation

rip 51

About Models of Computation
This was a core undergraduate Computer Science course on the foundations of computing. The questions it aimed to answer were: (1) What is computation? (2) What is computable? (3) What is computable given our limited resources? (4) How do we design, analyze, and evaluate algorithms? The course has now been renamed CS1010: Theory of Computation.

CS053 RELOADED
THE MATRIX IN COMPUTER SCIENCE

Home Resources Lectures Staff Homeworks Labs Documents Calendar Tips

The final exam will be held on December 16th, at 9am, in Solomon Center 903. Ari will be holding a final review session on December 15th from 7:30-9:30pm in CIT 368.

Make sure you are using the command `cs053_submit` to submit your auto-graded assignments.

Plag Highlights Out here or here.

What is The Matrix?

The aim of this course is to provide students interested in computer science an introduction to vectors and matrices and their use in CS applications. The course will be driven by applications from areas chosen from among: combinatorial optimization, computer vision, cryptography, game theory, graphics, information retrieval and web search, machine learning, and scientific visualization. For example, students will learn Google's PageRank method for ranking web pages. This course satisfies the linear algebra requirement for the Computer Science Sc.B and the Applied-Math/CS Sc.B.

There are no formal prerequisites, but students are expected to be comfortable with mathematics and programming.

Class meets in CIT 368 on Monday, Wednesday, and Friday, 3:50-4:20.
The course material can be found here.

Neutral



CS 106A

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.

Name	Due
Karel the Robot (INDIVIDUAL)	4/14
Consolation Prize (PAIR)	4/24
Hangman (PAIR)	5/1
Breakout (PAIR)	5/8
Melody Player + Image Algorithms (PAIR)	5/17
Critters (PAIR)	5/26
NameSurfer (PAIR)	6/5

SECTIONS

Our sections are 50-minute discussions led by section leaders each week in which the SL and students work together to solve problems. After Tue Apr 11 you can log in to our CS 198 site to see what section date/time/room you are in.

Name	Week
Section 1: Karel the Robot; Basic Java	4/12 - 4/14
Section 2: Java Control Statements, Parameters	4/19 - 4/21
Section 3: Strings/Characters; File Processing	4/26 - 4/28
Section 4: Graphics, Animation, and Fields	5/3 - 5/5
Section 5: Arrays, Images, Pixels	5/10 - 5/12
Section 6: Classes and Objects; Inheritance; Critters	5/17 - 5/19
Section 7: ArrayList and HashMap	5/24 - 5/26
Section 8: Graphical User Interfaces	5/31 - 6/2

Are you looking for the list of section leader names and section times? That information can be found on the Staff/SLs page.

Stereotypical



CS 106A

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.

Name	Due
Karel the Robot (INDIVIDUAL)	4/14
Consolation Prize (PAIR)	4/24
Hangman (PAIR)	5/1
Breakout (PAIR)	5/8
Melody Player + Image Algorithms (PAIR)	5/17
Critters (PAIR)	5/26
NameSurfer (PAIR)	6/5

SECTIONS

Our sections are 50-minute discussions led by section leaders each week in which the SL and students work together to solve problems. After Tue Apr 11 you can log in to our CS 198 site to see what section date/time/room you are in.

Name	Week
Section 1: Karel the Robot; Basic Java	4/12 - 4/14
Section 2: Java Control Statements, Parameters	4/19 - 4/21
Section 3: Strings/Characters; File Processing	4/26 - 4/28
Section 4: Graphics, Animation, and Fields	5/3 - 5/5
Section 5: Arrays, Images, Pixels	5/10 - 5/12
Section 6: Classes and Objects; Inheritance; Critters	5/17 - 5/19
Section 7: ArrayList and HashMap	5/24 - 5/26
Section 8: Graphical User Interfaces	5/31 - 6/2

Enrollment intentions

Sense of belonging



Anticipated Success

Experimental Measures

Technical self-confidence

Future CS intentions

Stereotype anxiety

 Neutral
 Stereotypical

Danaë Metaxa-Kakavouli, Kelly Wang, James A. Landay, and Jeff Hancock. 2018. Gender-Inclusive Design: Sense of Belonging and Bias in Web Interfaces. In Proceedings of CHI '18, 1–6. <https://doi.org/10.1145/3173574.3174188>

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

Severity Ratings Example

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.

Carl's Car Audio and Electronics -- Shopping Bag

what fits my car? [logout](#) [continue shopping](#)

Your vehicle: 1989 Tercel

To select a different vehicle, click 'What Fits My Car' above

PLEASE NOTE

Item #	Message
204EL570	This component is NOT recommended for your vehicle. We suggest removing this item (bolded in red below) from your cart. Please call us toll-free at 1-888-955-6000 and we'll be glad to provide further assistance.

Cart

Quantity	Item #	Remove?	Item Description	Price	Total
<input type="text" value="3"/>	#033DVM4800	<input type="checkbox"/>	Denon DVM-4800 DVD Video Changer	\$1,199.95	\$1,199.95
<input type="text" value="1"/>	#204EL570	<input type="checkbox"/>	Bazooka EL570 5"x7" 2-way speakers Vehicle: 1989 Toyota Tercel Liftback with equalizer or radio delete option (change)	\$79.95	\$79.95
<input type="text" value="h"/>	#158DSCP50	<input type="checkbox"/>	Special Sony DSC-P50 Digital Still Camera, 2.1 Mega Pixel Save \$100! Was \$399.95, Now: \$299.95	\$299.95	\$299.95
<input type="text" value="1"/>	#133DVDLA95	yes <input type="checkbox"/> no <input type="checkbox"/>	Panasonic DVD-LA95 Portable DVD-A/V Player with 9" Screen -- This item is temporarily out of stock --	\$999.95	\$999.95
<input type="text" value="1"/>	#170GXT160	<input type="checkbox"/>	Maxell GX-Silver T-160 VHS Video Tape	\$2.49	\$2.49
<input type="text" value="1"/>					

Total

Merchandise Total:			\$2,582.29
Standard Shipping Shipping Charge:			\$13.95
Alternative shipping options available before final checkout			
Order Total:			\$2,596.24

[Update Cart](#) [Clear Cart](#)

- 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

1. H4 Consistency and standards, severity 3
2. H4 - Which red refers to removing objects?
3. H5: Error prevention; the letter "h" is in Quantity, and no error is identified
4. H1: system status; "International Visitors, please click here" doesn't explain why they have to click there, or where it will take them
5. H4: consistency; in the remove column, it's either a toggle or a yes and a no check box
6. H4: consistency; "Total" means different things in different places
7. H5: you're allowed to order something that is out of stock? Severity: 2
8. H4, severity 2. Background color differs between blue and white
9. H4: message about one problem is at the top as opposed to under like out of stock one
10. [H5 Error Prevention][Level 2] Unclear that the message under Please Note is describing an error. Maybe make the error explanation text color red to match.
11. H11 - the title is not accessible to red-green color blindness
12. H9: gives the user the ability to not remove an out-of-stock item
13. H2: uses item numbers and codes that users don't understand [level 4]
14. H8: minimalist design not followed, lots of different colors without purpose
15. H4: Inconsistent coloring of boxes in the cart
16. h8: cluttered design with lots of extraneous information
17. H4, severity 3 Remove option is one option except for item 4 which is yes or no for some reason
18. H11 - severity 3 - coloring and contrast is not accessible for color blindness, the font is also really small
19. H8 - the design is extremely cluttered with all the information being all displayed in the same size and type (only diff is color), user needs to be able to identify the most important things, remove all irrelevant information, and highlight important information
20. H8: Aesthetic and minimalist design - too many different and vibrant colors for background and text
21. H11: accessible design [3-4] Illegible text, high saturation title and blue text, change colors
22. H8: Unnecessary information added that clusters the page, design is not minimalistic, conflicting colors
23. H5 Error Prevention the quantity is h instead of a number at row 3
24. H8: incredibly cluttered and confusing with different colors
25. H4: Inconsistent coloring of boxes in the cart
26. H4, severity 3 Remove option is one option except for item 4 which is yes or no for some reason
27. H7 severity 3, instructions are at the bottom rather than the top
28. H4 Consistency and Standards Severity 3: Uses links at the top to navigate (logout, continue shopping) but buttons for navigation elsewhere (update cart, clear cart), making navigation unnecessarily difficult
29. H4: Remove checkboxes are inconsistent, one has yes no
30. H4: What does red indicate? Used for discounts, stock, and item nums
31. H2: The user does not need to know the item number
32. H8 Aesthetic and Minimalist Design - Screen is cluttered and the colors aren't great - users don't care about item numbers, Price, and total seem to be the exact same thing, H6 [3] Recognition rather than recall, what does update cart mean
33. H4: checkbox for one product but not others
34. H8 - users don't need to know the item number, H4 - the color red is inconsistently used, H5 - this component is not recommended, H9 - DVD out of stock but still being charged, H2 - removing an item is confusing
35. H11: the colors aren't good for contrast (red text on a green background)
36. H5- add and clear cart buttons look the same
37. H4: remove boxes either have one checkbox or a yes/no checkboxes
38. H8 Aesthetic and Minimalist Design
39. H4, 2. Background color differs. H8, 4 designs not minimal and have a lot of distracting info
40. H1 - The error message is very hard to read
41. H5, H10 (?): issue with the second item #204... is explained in a message at the top of the screen but has to cross-check the item number with the list of items below... not immediately clear what the message is referring to and why #204... is read
42. H9: Error message, which indicates a product is not compatible with a car, is not easy to spot and requires a multiple step solution
43. H5 [Severity 4]: No error message given for the quantity of the digital still camera being "h". This is a fatal error because we have no idea how many the user actually meant to buy. Inform the user of this incorrect input or only allow them to input numerals.
44. H9: error message doesn't contain ways to resolve the error
45. H5: the component that is not recommended is written in black font and doesn't alert the user, should stand out more to alert the user before purchasing
46. H4: Use of the color red is inconsistent, means different things in different contexts
47. H1: Green decorative background of title text can confuse users into thinking that they have already completed a process, since green is associated with completion
48. H3 [Severity 4]: I can't see a checkout button. Even if it should not be accessible until I resolve all the errors, it should be greyed out or similar so I don't feel unable to proceed.
49. H4: color red is used for a lot of different things (errors, ads, etc.)
50. H2: don't need to know the item # (could also be error prevention)
51. H9: out of stock but still being charged
52. H9: there is an h in quantity but no indication
53. H7/10: too many steps to update cart
54. H8: blue text on blue background

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

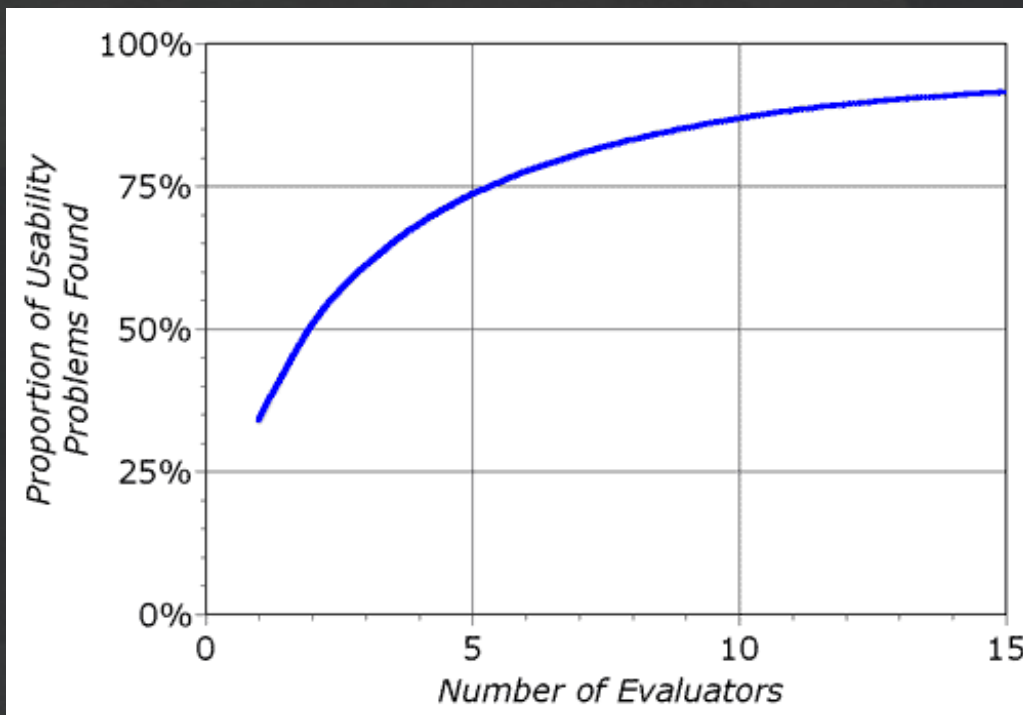
- H9: Help users recognize, diagnose, & recover from errors
 - “Please Note” error message refers to item bolded in red but multiple items bolded in red [2]
 - Last row has quantity 1 but no other info, user unsure how to fix [1]
- H10: Help & Documentation
 - Instructions for use not immediately visible [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

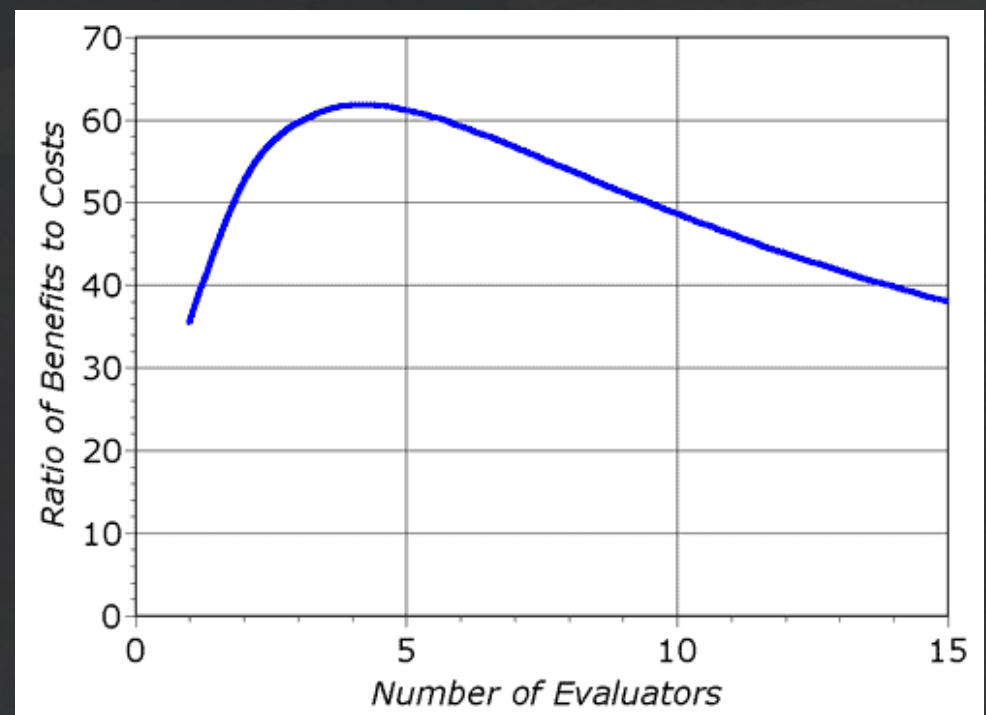
- H4 (consistency): The use of red to indicate errors, out of stock items, and sales/savings [26]
- H4 (consistency): two of the check boxes have yes/no next to them but none of the others do. Error prevention? [22]
- H9 (Aesthetic): "#" is not needed [2]
- H5 [Error Prevention]: "h" in the quantity box. [11]

Decreasing Returns

problems found



benefits / cost



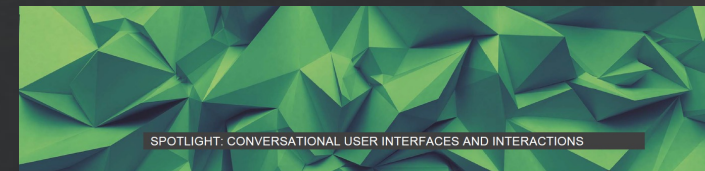
* Caveat: graphs for a specific example

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.



SPOTLIGHT: CONVERSATIONAL USER INTERFACES AND INTERACTIONS

Evaluating Speech-Based Smart Devices Using New Usability Heuristics

Zhuxiaona Wei
depleaming.ai
James A. Landay
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 almost three times faster than typing a short message.¹ Thanks to recent breakthroughs in speech and language technologies, speech user interfaces (SUIs) have improved rapidly, and voice-enabled devices are now common. Baidu's Deep Speech 2 system, for example, can recognize spoken words with human-level accuracy.²

Nevertheless, designing good SUIs remains challenging.³ The state of an SUI is often opaque to users, leading to more user errors compared to graphical user interfaces (GUIs).⁴ Unfortunately, simply transforming GUIs into speech interfaces does not work well.⁵ Although researchers have been working on SUI technology for three decades, much useful knowledge is in older papers and not easily accessible to designers. Moreover, the knowledge has not been updated to reflect recent improvements in speech-recognition accuracy. Consequently, those new to SUI design often feel lost.⁶

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 had a group of usability experts—half of whom specialized in SUIs—use them 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/articles/ten-usability-heuristics/). Although these heuristics are most often applied to GUIs, he and his colleagues also used them to evaluate a telephone voice-response system.⁷ However, the user input and system output options for the system were quite limited.

IEEE Pervasive Computing
April–June 2018

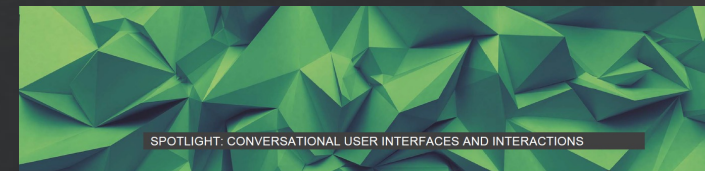
84

Published by the IEEE Computer Society
1536-1268/18/0333-00-02018 IEEE

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>



Evaluating Speech-Based Smart Devices Using New Usability Heuristics

Zhuxiaona Wei
depleaming.ai
James A. Landay
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 almost three times faster than typing a short message.¹ Thanks to recent breakthroughs in speech and language technologies, speech user interfaces (SUIs) have improved rapidly, and voice-enabled devices are now common. Baidu's Deep Speech 2 system, for example, can recognize spoken words with human-level accuracy.²

Nevertheless, designing good SUIs remains challenging.³ The state of an SUI is often opaque to users, leading to more user errors compared to graphical user interfaces (GUIs).⁴ Unfortunately, simply transforming GUIs into speech interfaces does not work well.⁵ Although researchers have been working on SUI technology for three decades, much useful knowledge is in older papers and not easily accessible to designers. Moreover, the knowledge has not been updated to reflect recent improvements in speech-recognition accuracy. Consequently, those new to SUI design often feel lost.⁶

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 had a group of usability experts—half of whom specialized in SUIs—use them 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/articles/ten-usability-heuristics). Although these heuristics are most often applied to GUIs, he and his colleagues also used them to evaluate a telephone voice-response system.⁷ However, the user input and system output options for the system were quite limited.

IEEE Pervasive Computing
April–June 2018

84

Published by the IEEE Computer Society
1536-1268/18/0333-00-02018 IEEE

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** be 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%

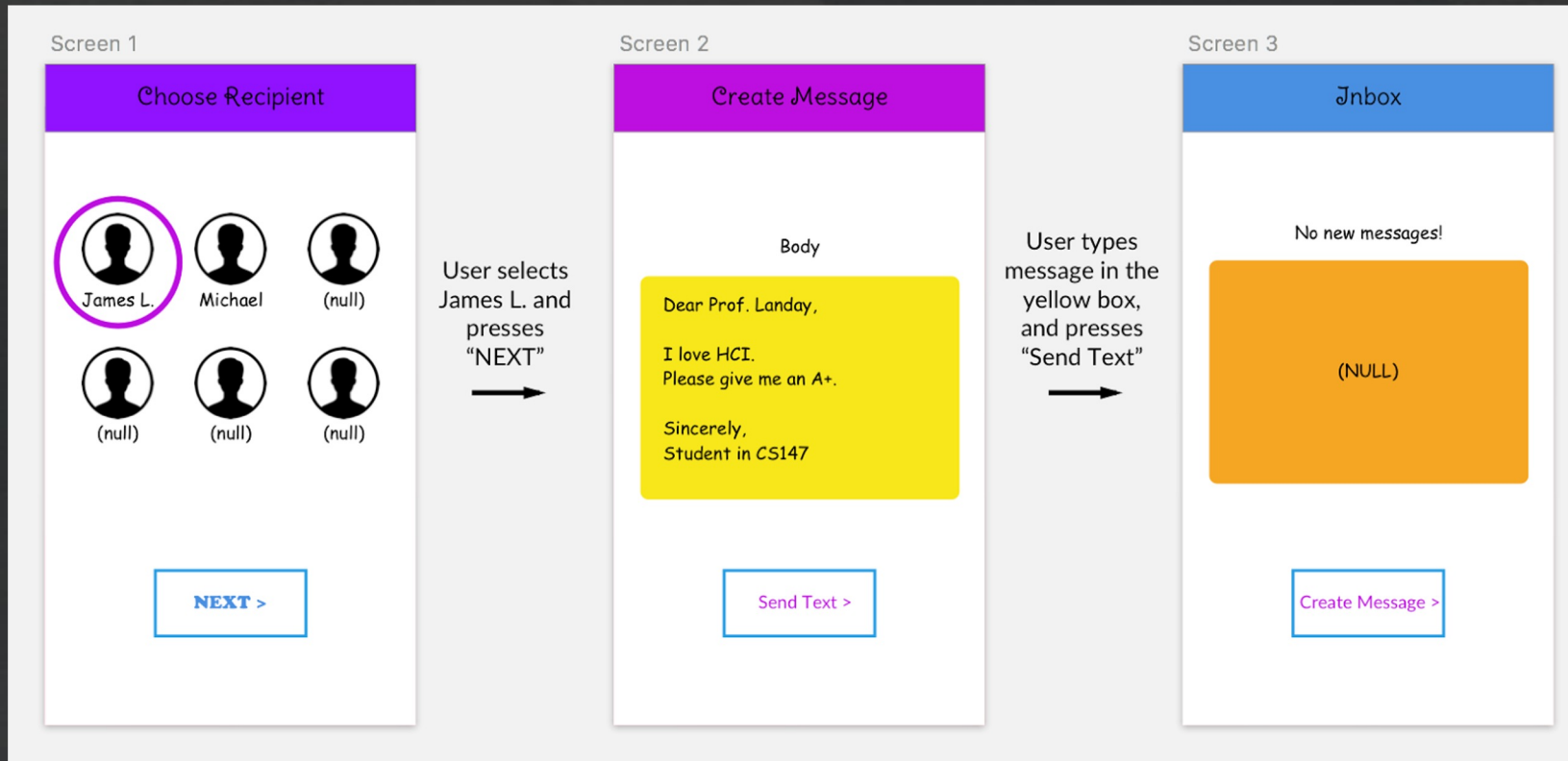
TEAM BREAK

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?

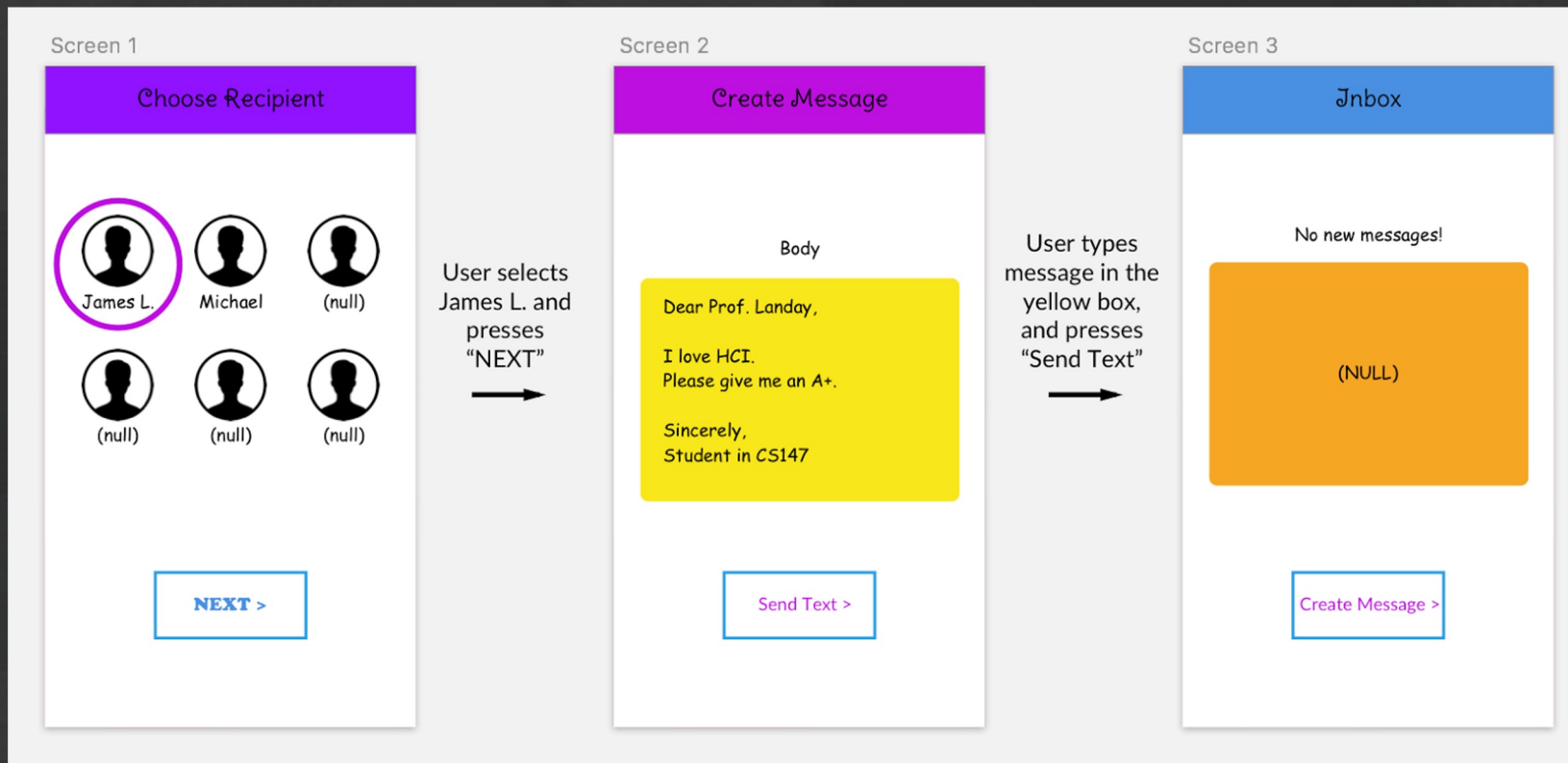


<https://bit.ly/HE-cs147-23au>



[5 min]: Share with your group (3-4)

What violations did you all find? Which ones did only one of you notice?

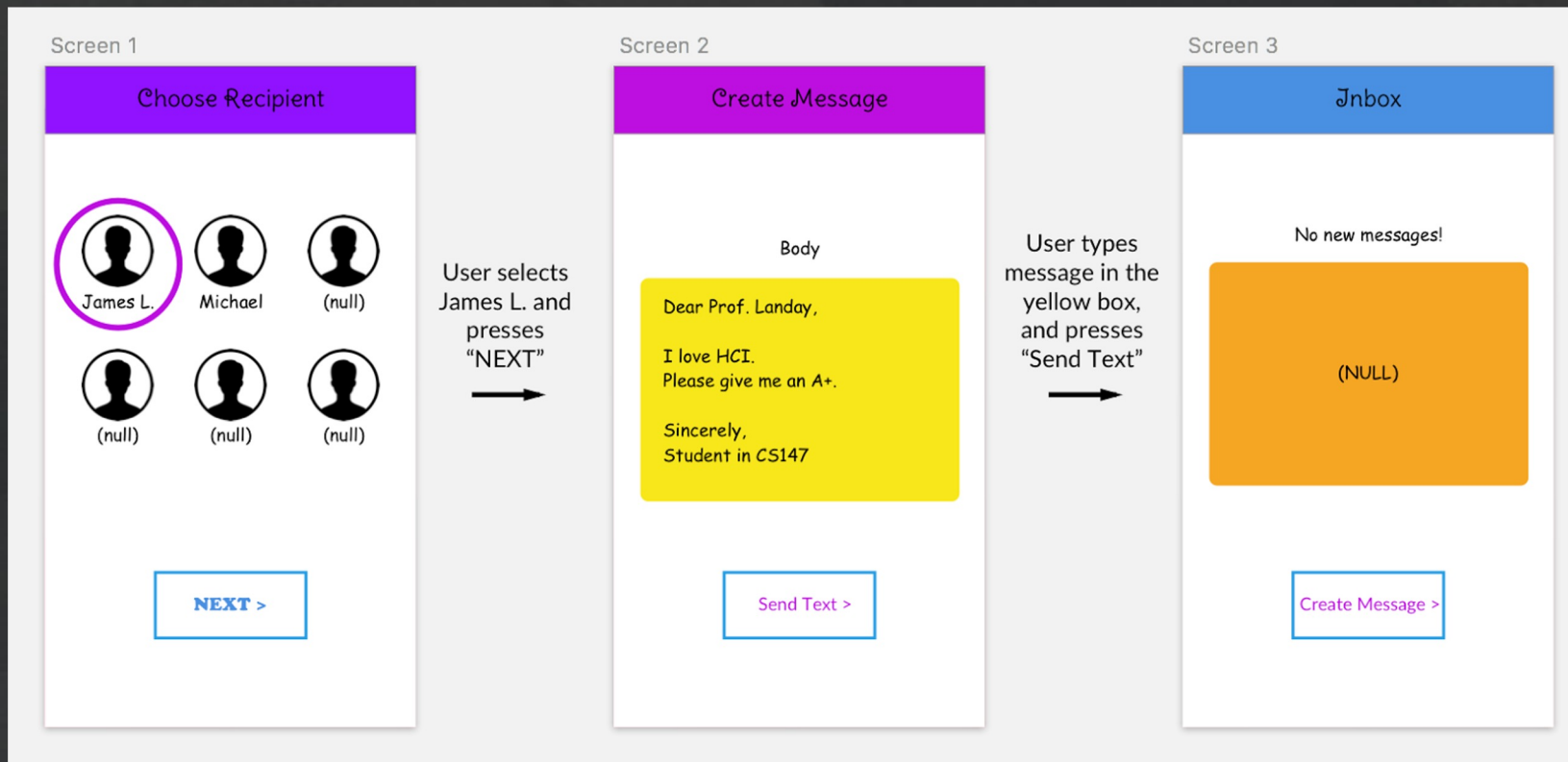


<https://bit.ly/HE-cs147-23au>



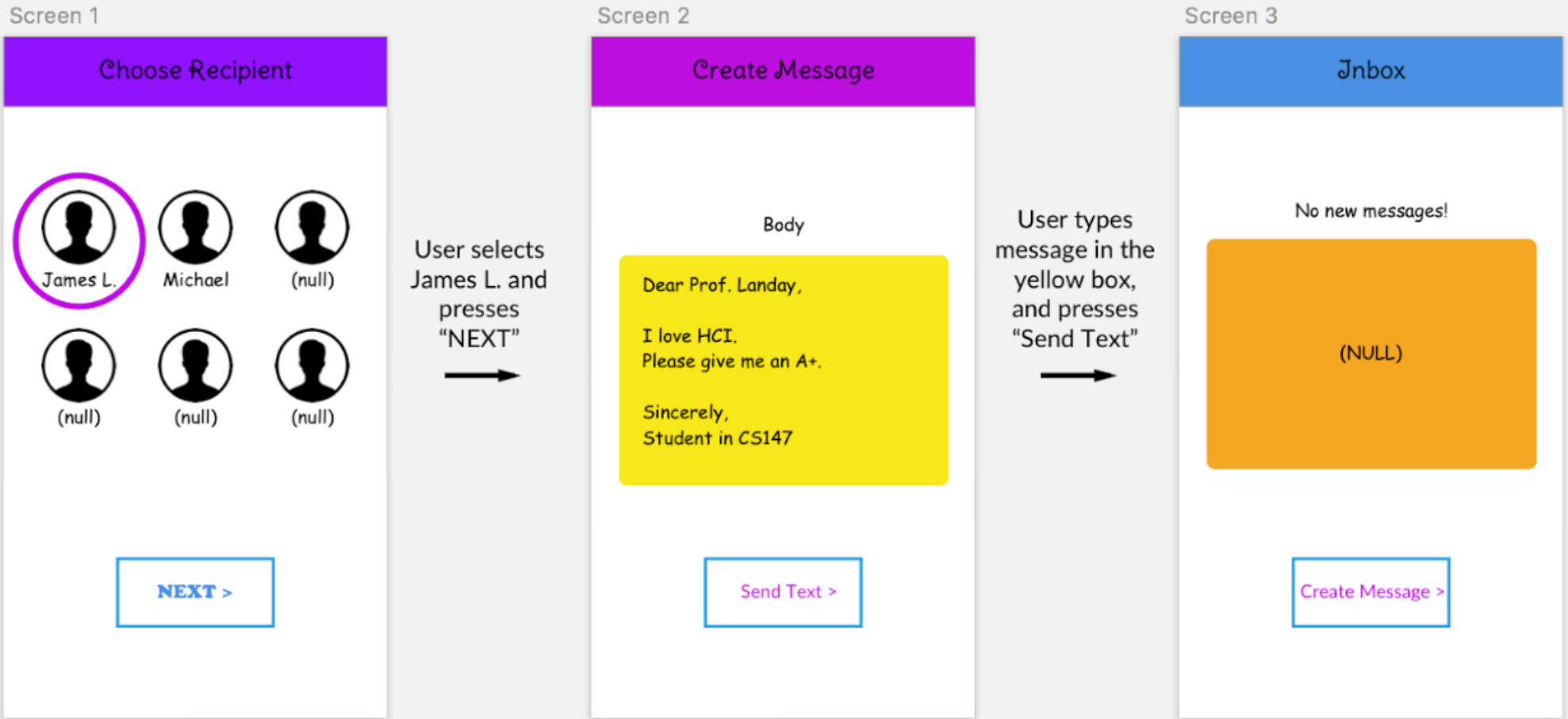
[5 min]: Share with the class

Raise your hand or post a screenshot with your description on Slack.



<https://bit.ly/HE-cs147-23au>

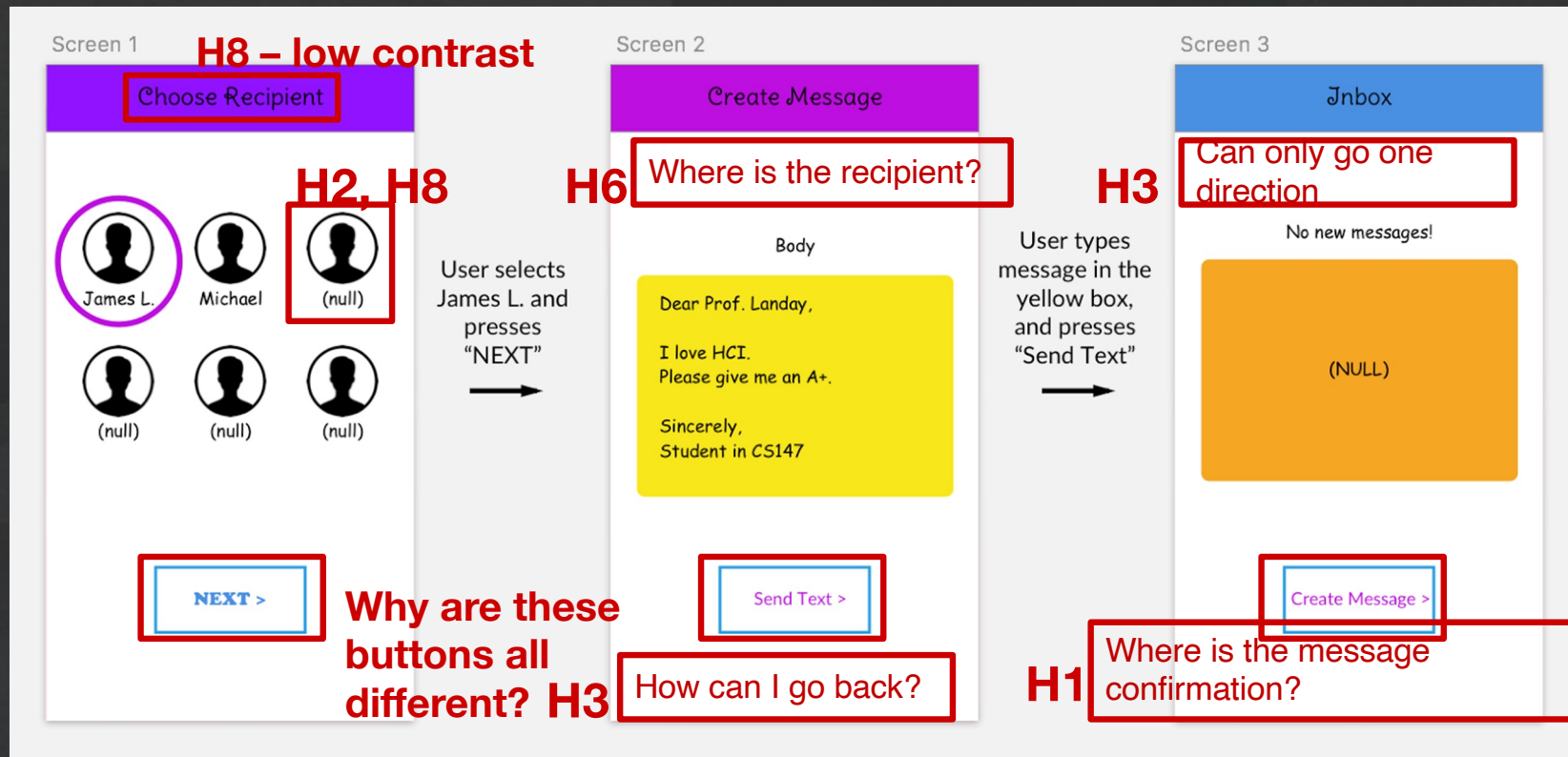




Find, label, & describe 8-10 Heuristic Violations

[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).



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 mislicked 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/>
- Accessibility
 - [Accessibility - Foundations - Human Interface Guidelines - Design - Apple Developer](#)

Exit Ticket

<http://bit.ly/CS147-2023au-exit-ticket-6-937>

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