dt+UX DESIGN THINKING FOR USER EXPERIENCE DESIGN + PROTOTYPING + EVALUATION

Heuristic Evaluation

刘晋明
Prof. James A. Landay
Computer Science Department
Stanford University

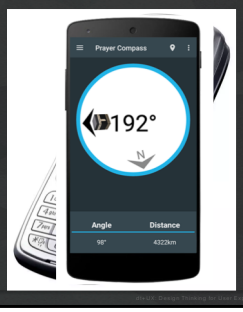
Autumn 2018
October 31, 2018

Hall of Fame or Shame?

LG F7100
courtesy of Genevieve Bell, Intel

Launched in 2004 in UAE, Saudia Arabia, North Africa, India, Malaysia

Hall of Fame!



LG F7100
courtesy of Genevieve Bell, Intel

Launched in 2004 in UAE, Saudia Arabia, North Africa, India, Malaysia

Good
- targeted at Muslim audience
- need to pray 5x/day pointing towards Mecca

dt+UX DESIGN THINKING FOR USER EXPERIENCE DESIGN + PROTOTYPING + EVALUATION

Heuristic Evaluation

刘晋明
Prof. James A. Landay
Computer Science Department
Stanford University

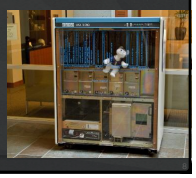
Autumn 2018
October 31, 2018

Outline

- Wizard of Oz
- Heuristic Evaluation Overview
- The Heuristics
- Team Break
- Exercise

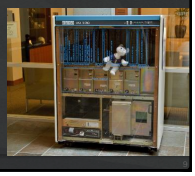
Wizard of Oz Technique

- Faking the interaction. Comes from?
  - the film "The Wizard of OZ"
    - "the man behind the curtain"
- Long tradition in computer industry
  - e.g., prototype of a PC w/ a DEC VAX behind the curtain



Wizard of Oz Technique

- Faking the interaction. Comes from?
  - the film "The Wizard of OZ"
    - "the man behind the curtain"
- Long tradition in computer industry
  - e.g., prototype of a PC w/ a DEC VAX behind the curtain
- Much more important for hard to implement features



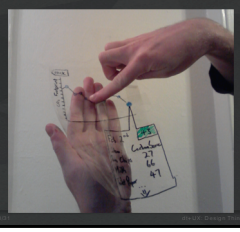
Wizard of Oz Technique

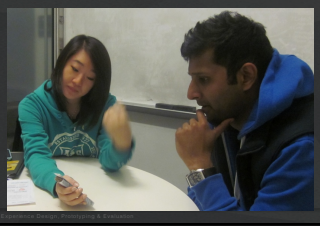
Carbon Shopper

- Much more important for hard to implement features

Evaluation

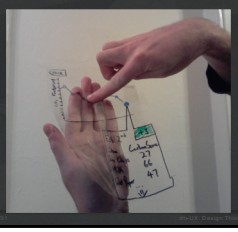
- About figuring out how to improve design
- Issues with lo-fi tests?





Evaluation

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



Not realistic
– visuals & performance

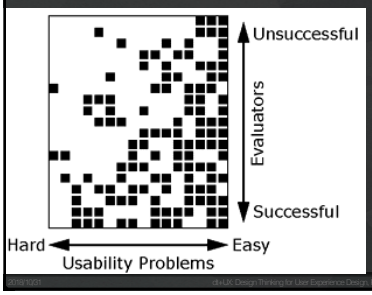
Not on actual interface
– can't test alone

Need participants
– can be hard to find repeatedly

Heuristic Evaluation

- Developed by Jakob Nielsen
- Helps find usability problems in a UI design
- 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?



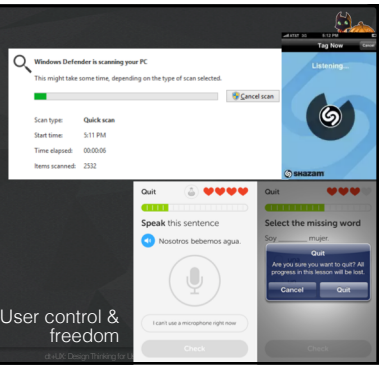
- Every evaluator doesn't find every problem
- Good evaluators find both easy & hard ones

Heuristics

H1: Visibility of system status

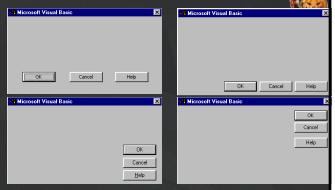
H2: Match between system & real world

H3: User control & freedom



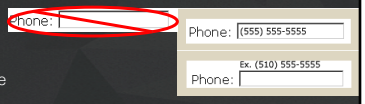
Heuristics (cont.)

H4: Consistency & standards



H5: Error prevention

H6: Recognition rather than recall

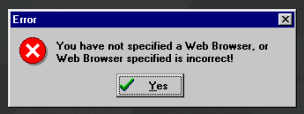


H7: Flexibility and efficiency of use

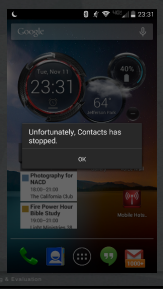
H8: Aesthetic & minimalist design

Heuristics (cont.)

bad

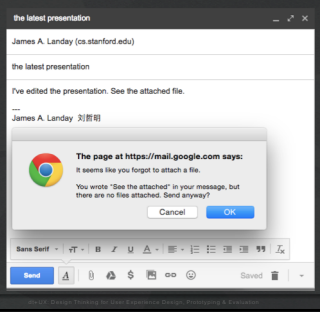


H9: Help users recognize, diagnose, & recover from errors



Heuristics (cont.)

good

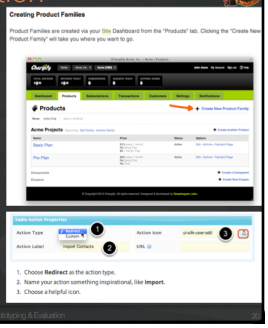


Good Error Messages

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

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

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

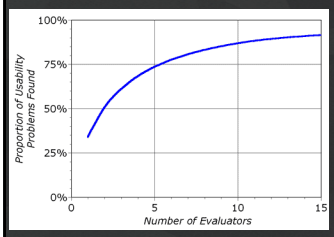
1. [H4 Consistency & Standards] [Severity 3]

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.

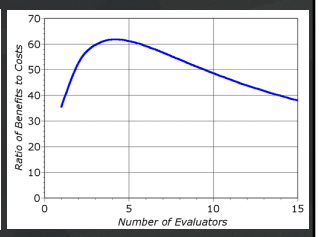
Fix: Use "Save" everywhere in the application.

Decreasing Returns

problems found



benefits / cost



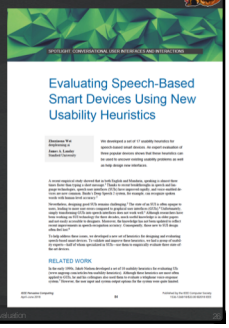
* Caveat: graphs for a specific example

Heuristic Evaluatoin 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

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

Christiane Wei Phillipsborn, Jan-Ole Jonas, Sebastian Gerstenberg, Jens Lüssem, Matthias Wenzel, Rainer Malaka

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 magazine article claimed that in both English and Mandarin, speaking is almost three times faster than typing a short message.[1] Supply to recent breakthroughs in speech and language technologies, speech now outclasses QWERTY-based typing, and voice enabled interfaces are rapidly displacing Graphical User Interfaces (GUIs). For example, voice recognition systems with human-level accuracy.[2]

Nevertheless, designing good VUIs remains challenging.[3] The state of art ICT is often opaque to users, and it is difficult to predict what a system can understand and what it cannot. Moreover, speech input/output (VUIs) are special, unlike a dare and rude tool.[4] Although extensive have been work evolving on VUI technology for three decades, much works knowledge is in other papers. Moreover, most of the work is in the form of case studies and anecdotal evidence, and there are recent experiments in speech recognition systems. Consequently, there are in VUI design guidelines.

To help address these issues, we developed a 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 them specialized in VUIs—use them to empirically evaluate three state-of-the-art devices.

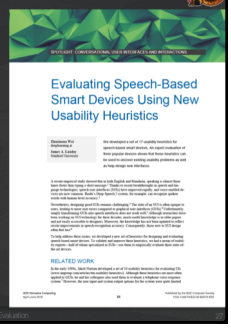
RELATED WORK

In the early 1990s, Jakob Nielsen developed a set of 10 usability heuristics for evaluating GUIs (www.nngroup.com/articles/ten-usability-heuristics). Although these heuristics are more often applied to web sites, they are also used to evaluate other types of interfaces, including voice systems. Moreover, the user input and system output options for the system were quite limited.

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


SPOTLIGHT: CONVERSATIONAL USER INTERFACES AND INTERACTIONS

Evaluating Speech-Based Smart Devices Using New Usability Heuristics

Dissiemon Wei
Implementing a
Jamie 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 the ubiquitous adoption of smartphones, speech is now a viable input method for many mobile devices even pervasive contexts. Both Chinese Speech Corpora, for example, can recognize spoken Mandarin with 97% accuracy.[2]

Nevertheless, designing good SUIs remains challenging.[3] The state of an SUI's often opaque to users, leading to more error issues compared to graphical user interfaces (GUIs).[4] Unfortunately, there are few guidelines for designers to follow. Although Apple, Google, and Microsoft have been working on SUI technology for three decades, much needed knowledge is in other papers and proprietary documents. Moreover, most of the SUI design guidelines are based on early voice-only interactions in speech recognition systems. Consequently, there are few SUI design ideas for two.[5]

To help address these issues, we developed a new set of heuristics for designing and evaluating smart speaker devices. We validate and improve these heuristics, we find 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 GUIs (www.nngroup.com/articles/ten-usability-heuristics). Although these heuristics are most often used for GUIs, they have been used for other types of interfaces, including voice-only interfaces.[6] However, the non-expert and system output options for the system were quite limited.

IEEE Pervasive Computing
April–June 2018

38

Published by the IEEE Computer Society
Digital Object Identifier 10.1109/MPRV.2018.

2018/10/31 dt+UX: Design Thinking for User Experience Design, Prototyping & Evaluation 22

Administrivia

- Free Figma pro licenses if you use your Stanford emails to verify (instructions on Piazza)
- Questions on medium-fi prototype assignment?

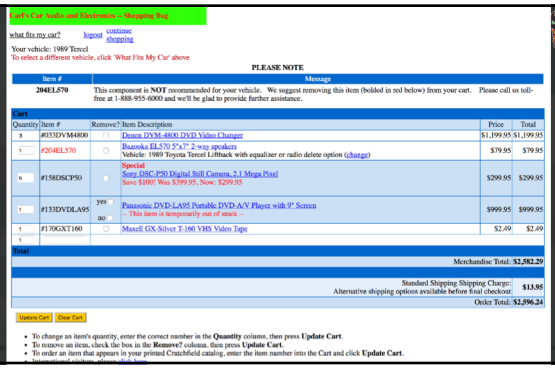
2018/10/31 dt+UX: Design Thinking for User Experience Design, Prototyping & Evaluation 23

TEAM BREAK

2018/10/31 dt+UX: Design Thinking for User Experience Design, Prototyping & Evaluation 24

E X E R C I S E

2018/10/31 dt+UX: Design Thinking for User Experience Design, Prototyping & Evaluation 25


Car-Tech 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
3	#033DVM4800		Denon DVM-4800 DVD Video Changer	$1,199.95	$1,199.95
1	#204EL570		Bazooka EL570 5"x7" 2-way speakers Vehicle: 1989 Toyota Tercel Liftback with equalizer or radio delete option (change)	$79.95	$79.95
1	#158DSCP50		Special Sony DSC-P50 Digital Still Camera, 2.1 Mega Pixel Save: $1001 Was $399.95, Now: $299.95	$299.95	$299.95
1	#133DVDLA95	yes no	Panasonic DVD-LA95 Portable DVD A/V Player with 9" Screen -- This item is temporarily out of stock .	$999.95	$999.95
1	#170GXT160		Maxell GX-Silver T-160 VHS Video Tape	$2.49	$2.49
1					
Total					
			Merchandise Total: **$2,582.29**		
			Standard Shipping Charge: Alternative shipping options available before final checkout:	**$13.95**	
			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**.

2018/10/31 dt+UX: Design Thinking for User Experience Design, Prototyping & Evaluation 26


Find 12-15 Heuristic Violations

2018/10/31 dt+UX: Design Thinking for User Experience Design, Prototyping & Evaluation 27

Problems Found

1. ?

Problems Found Last Year

1. H3 – no purchase button [100]
2. H4 – remove column has check boxes and then one entry w/ yes/no [100]
3. H5 – illegal input (text) allowed in quantity field [90]
4. H1 – not clear who is logged in [10]
5. H2 – "what fits my car" is not a good term that people would know [20]
6. H5 – user can add "out of stock" items [80]

Problems Found Two Years Ago

1. H5 Error Prevention
   allows non-numeric data in the quantity field. fix: don't allow it. [90]
2. H5 Error Prevention
   quantity field doesn't multiply by the price to give a correct total. fix: make it work. [55]
3. H10 Help & Documentation
   link for international visitors hidden at bottom & may not be readable by non-english speakers. fix: move up to prominent location & include flags? [30]
4. H5 Error Prevention
   "Remove item bolded in red", but red used for multiple purposes. Fix: get rid of ads in the checkout! More direct way to remove out of stock item or not even let me add a item that is out of stock. [100]
5. H5 Error Prevention
   No way to check out. [110]

Problems Found Three Years Ago

1. H4 Consistency
   remove column, 4th item is different w/ checkboxes. [150]
2. H9 Error prevention
   non-numeric data in the quantity. Do not allow. [125]
3. H2 Match between system & real world
   vehicle selection link not language I'd expect [100]
4. H1 Visibility of System Status
   unclear which item to remove based on error message ("red/bold"). [150]

Further Reading

Heuristic Evaluation

- Longer lecture
  - https://drive.google.com/file/d/0BweiB6wu4sBaN2tfZGxKb2tuOTg/view
- Books
  - Usability Engineering, by Nielsen, 1994
- Web site
  - http://www.nngroup.com/articles/

Next Time

- Lecture
  - Human Abilities
- Readings
  - Pages 66-99 of "Cognitive Aspects in Interaction Design", from *Interaction Design*, 3rd Edition by Rogers, Sharp, & Preece
- Next assignments
  - Individual Heuristic Evaluation
  - Group Heuristic Evaluation (in studio)