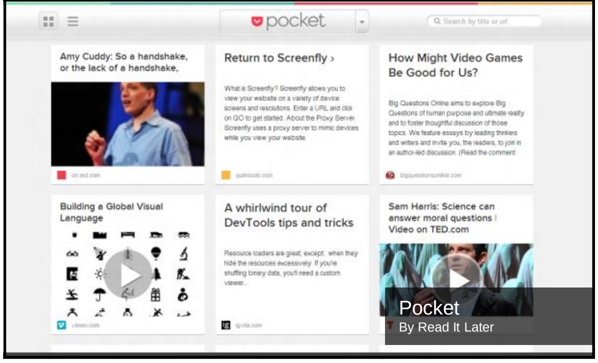
USER INTERFACE DESIGN + PROTOTYPING + EVALUATION

Heuristic Evaluation

Prof. James A. Landay
University of Washington

Autumn 2012

Hall of Fame or Shame?



Amy Cuddy: So a handshake, or the lack of a handshake,

on.ted.com

Return to Screenfly ›

What is Screenfly? Screenfly allows you to view your website on a variety of device screens and resolutions. Enter a URL and click on GO to get started. About the Proxy Server Screenfly uses a proxy server to mimic devices while you view your website.

quirktools.com

How Might Video Games Be Good for Us?

Big Questions Online aims to explore Big Questions of human purpose and ultimate reality and to foster thoughtful discussion of those topics. We feature essays by leading thinkers and writers and invite you, the readers, to join in an author-led discussion. (Read the comment

bigquestionsonline.com

Building a Global Visual Language

vimeo.com

A whirlwind tour of DevTools tips and tricks

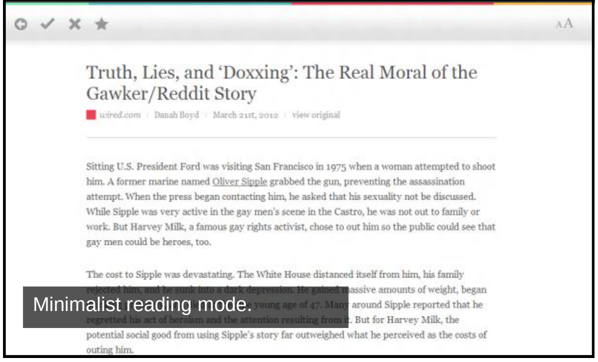
Resource loaders are great, except. when they hide the resources excessively. If you're shuffling binary data, you'll need a custom viewer...

igvita.com

Sam Harris: Science can answer moral questions | Video on TED.com

Pocket
By Read It Later

Hall of Fame or Shame?



Truth, Lies, and 'Doxxing': The Real Moral of the Gawker/Reddit Story

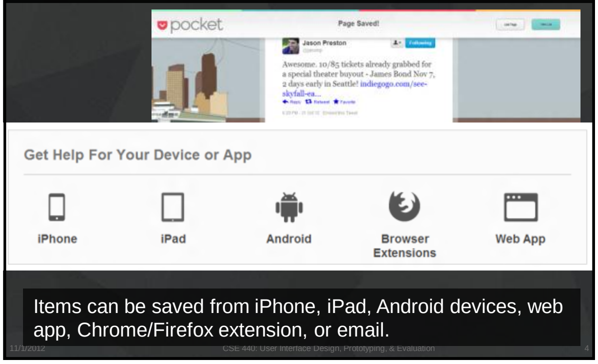
wired.com | Danah Boyd | March 21st, 2012 | view original

Sitting U.S. President Ford was visiting San Francisco in 1975 when a woman attempted to shoot him. A former marine named Oliver Sipple grabbed the gun, preventing the assassination attempt. When the press began contacting him, he asked that his sexuality not be discussed. While Sipple was very active in the gay men's scene in the Castro, he was not out to family or work. But Harvey Milk, a famous gay rights activist, chose to out him so the public could see that gay men could be heroes, too.

The cost to Sipple was devastating. The White House distanced itself from him, his family rejected him, and he sunk into a dark depression. He gained massive amounts of weight, began [illegible] young age of 47. Many around Sipple reported that he regretted his act of heroism and the attention resulting from it. But for Harvey Milk, the potential social good from using Sipple's story far outweighed what he perceived as the costs of outing him.

Minimalist reading mode.

Hall of Fame or Shame?



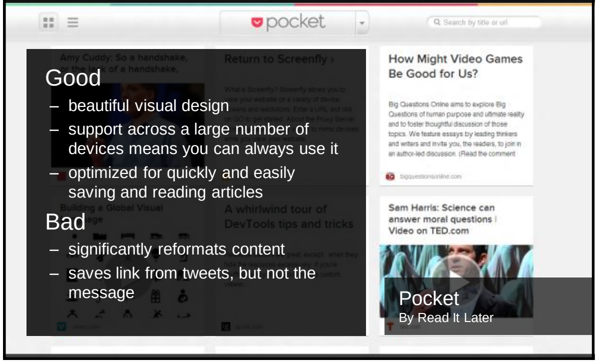
Awesome. 10/85 tickets already grabbed for a special theater buyout - James Bond Nov 7, 2 days early in Seattle! indiegogo.com/see-skyfall-ea...

Get Help For Your Device or App

Items can be saved from iPhone, iPad, Android devices, web app, Chrome/Firefox extension, or email.

11/1/2012 CSE 440: User Interface Design, Prototyping, & Evaluation 4

Hall of Fame!



How Might Video Games Be Good for Us?

Big Questions Online aims to explore Big Questions of human purpose and ultimate reality and to foster thoughtful discussion of those topics. We feature essays by leading thinkers and writers and invite you, the readers, to join in an author-led discussion. (Read the comment

bigquestionsonline.com

Sam Harris: Science can answer moral questions | Video on TED.com

Good
- beautiful visual design
- support across a large number of devices means you can always use it
- optimized for quickly and easily saving and reading articles

Bad
- significantly reformats content
- saves link from tweets, but not the message

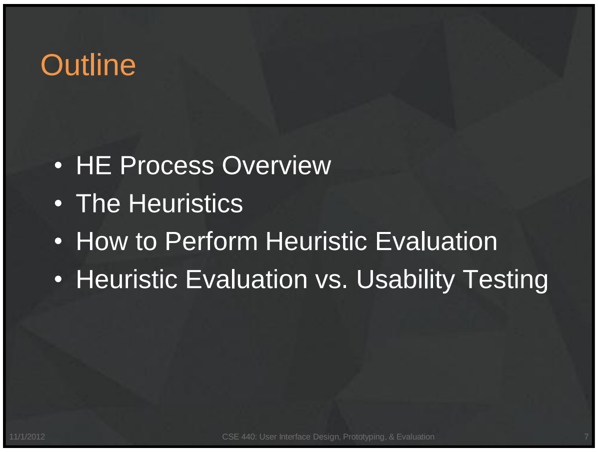
Pocket
By Read It Later

USER INTERFACE DESIGN + PROTOTYPING + EVALUATION

Heuristic Evaluation

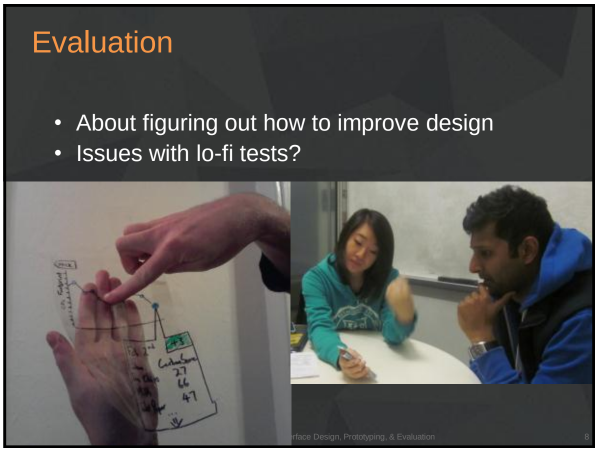
Prof. James A. Landay
University of Washington

Autumn 2012



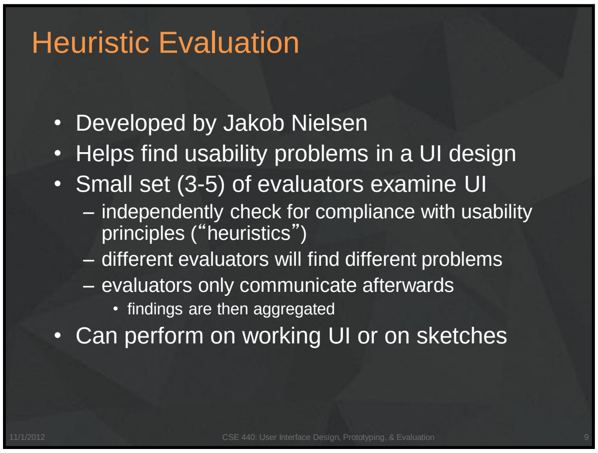
Outline

- HE Process Overview
- The Heuristics
- How to Perform Heuristic Evaluation
- Heuristic Evaluation vs. Usability Testing



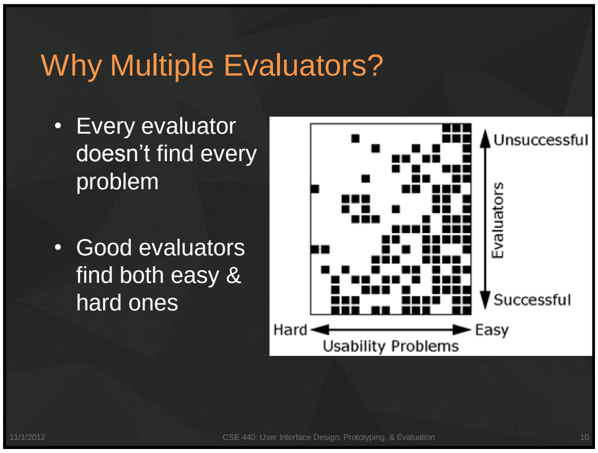
Evaluation

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



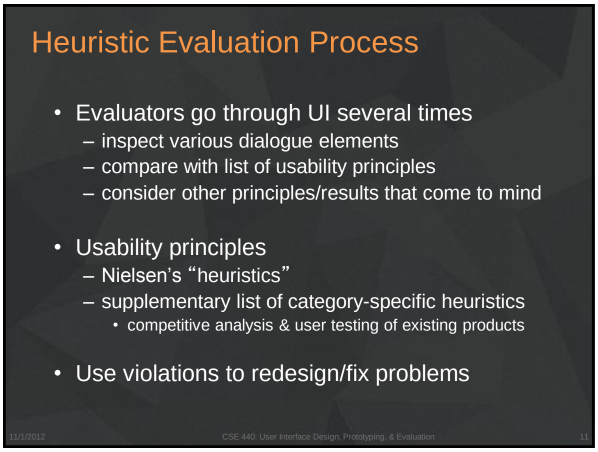
Heuristic Evaluation

- 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")
  - different evaluators will find different problems
  - evaluators only communicate afterwards
    - findings are then aggregated
- Can perform on working UI or on sketches



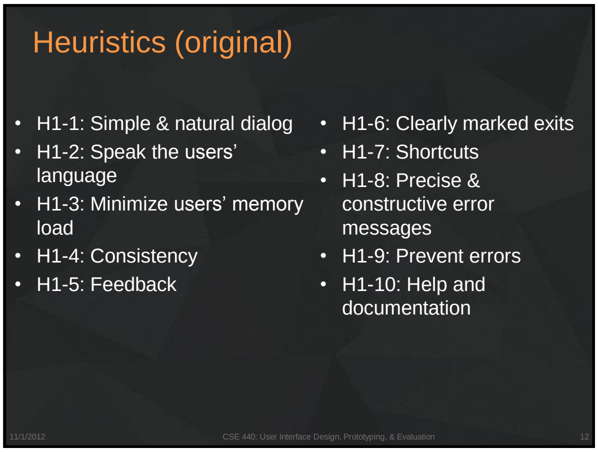
Why Multiple Evaluators?

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



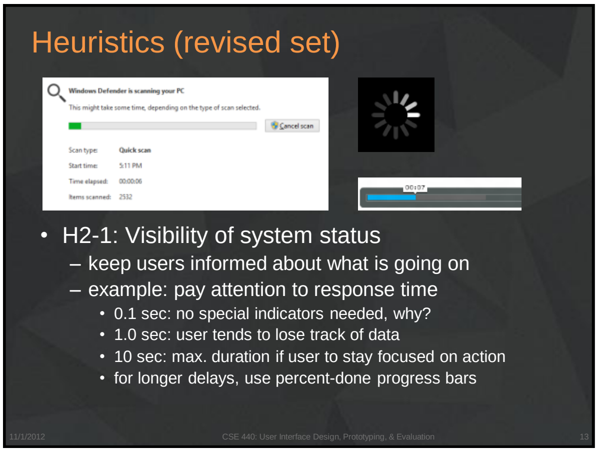
Heuristic Evaluation Process

- Evaluators go through UI several times
  - inspect various dialogue elements
  - compare with list of usability principles
  - consider other principles/results that come to mind
- Usability principles
  - Nielsen's "heuristics"
  - supplementary list of category-specific heuristics
    - competitive analysis & user testing of existing products
- Use violations to redesign/fix problems



Heuristics (original)

- H1-1: Simple & natural dialog
- H1-2: Speak the users' language
- H1-3: Minimize users' memory load
- H1-4: Consistency
- H1-5: Feedback
- H1-6: Clearly marked exits
- H1-7: Shortcuts
- H1-8: Precise & constructive error messages
- H1-9: Prevent errors
- H1-10: Help and documentation

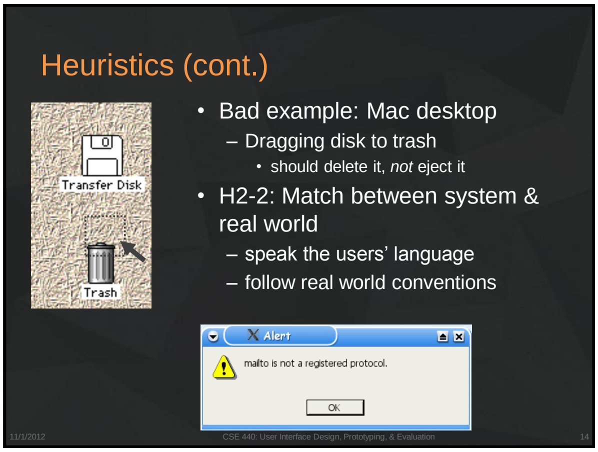


Heuristics (revised set)

Windows Defender is scanning your PC
This might take some time, depending on the type of scan selected.

Scan type: Quick scan
Start time: 5:11 PM
Time elapsed: 00:00:06
Items scanned: 2532

- H2-1: Visibility of system status
  - keep users informed about what is going on
  - example: pay attention to response time
    - 0.1 sec: no special indicators needed, why?
    - 1.0 sec: user tends to lose track of data
    - 10 sec: max. duration if user to stay focused on action
    - for longer delays, use percent-done progress bars

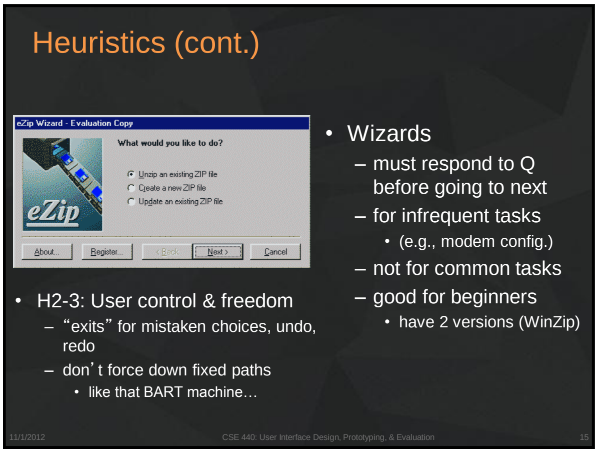


Heuristics (cont.)

- Bad example: Mac desktop
  - Dragging disk to trash
    - should delete it, *not* eject it
- H2-2: Match between system & real world
  - speak the users' language
  - follow real world conventions

Alert

mailto is not a registered protocol.



Heuristics (cont.)

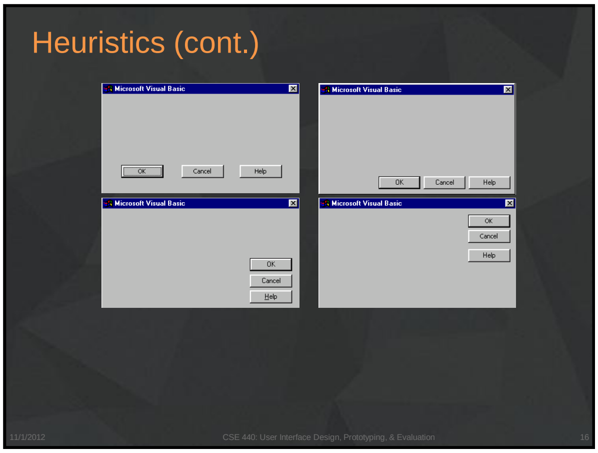
eZip Wizard - Evaluation Copy

What would you like to do?

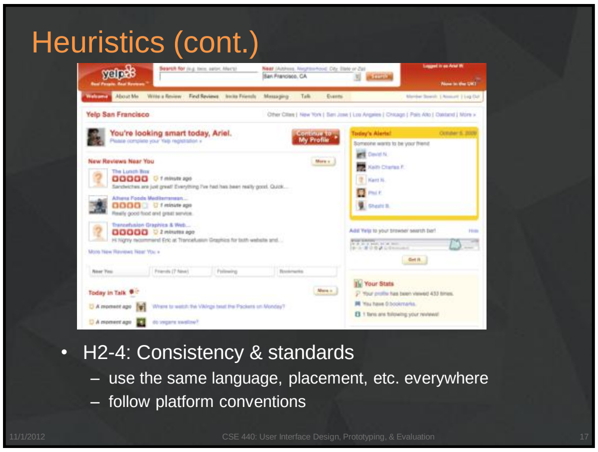
● Unzip an existing ZIP file
○ Create a new ZIP file
○ Update an existing ZIP file

- Wizards
  - must respond to Q before going to next
  - for infrequent tasks
    - (e.g., modem config.)
  - not for common tasks
  - good for beginners
    - have 2 versions (WinZip)

- H2-3: User control & freedom
  - "exits" for mistaken choices, undo, redo
  - don't force down fixed paths
    - like that BART machine…



Heuristics (cont.)



Heuristics (cont.)

Yelp San Francisco

You're looking smart today, Ariel.

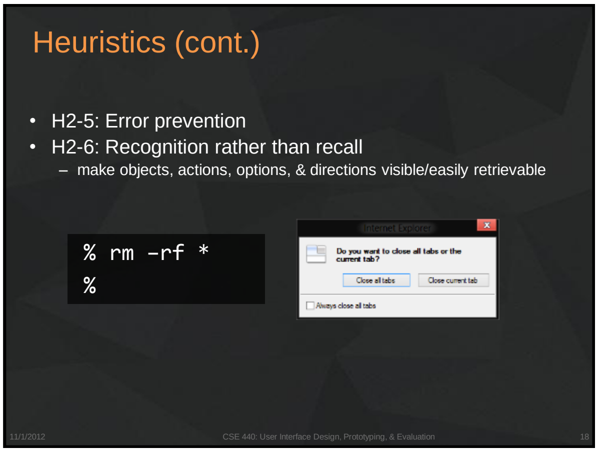
New Reviews Near You

Today's Alerts!

Today in Talk

Your Stats

- H2-4: Consistency & standards
  - use the same language, placement, etc. everywhere
  - follow platform conventions



Heuristics (cont.)

- H2-5: Error prevention
- H2-6: Recognition rather than recall
  - make objects, actions, options, & directions visible/easily retrievable

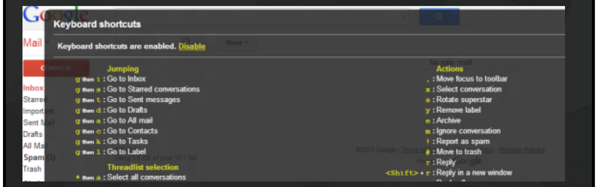
```
% rm -rf *
%
```

Internet Explorer

Do you want to close all tabs or the current tab?

□ Always close all tabs

Heuristics (cont.)



Keyboard shortcuts

Keyboard shortcuts are enabled. Disable

Jumping

- g then i : Go to Inbox
- g then s : Go to Starred conversations
- g then t : Go to Sent messages
- g then d : Go to Drafts
- g then a : Go to All mail
- g then c : Go to Contacts
- g then k : Go to Tasks
- g then l : Go to Label

Threadlist selection

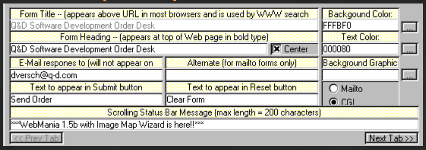
- * then a : Select all conversations

Actions

- , : Move focus to toolbar
- x : Select conversation
- s : Rotate superstar
- y : Remove label
- e : Archive
- m : Ignore conversation
- ! : Report as spam
- # : Move to trash
- r : Reply
- <Shift> + r : Reply in a new window

- H2-7: Flexibility and efficiency of use
  - accelerators for experts (e.g., gestures, kb shortcuts)
  - allow users to tailor frequent actions (e.g., macros)

Heuristics (cont.)



Form Title -- (appears above URL in most browsers and is used by WWW search
Q&D Software Development Order Desk

Backgound Color:
FFFBF0

Form Heading -- (appears at top of Web page in bold type)
Q&D Software Development Order Desk ☒ Center

Text Color:
000080

E-Mail respones to (will not appear on
dversch@q-d.com

Alternate (for mailto forms only)

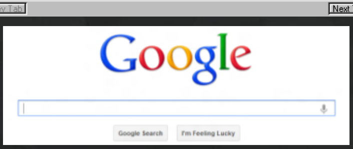
Background Graphic

Text to appear in Submit button
Send Order

Text to appear in Reset button
Clear Form

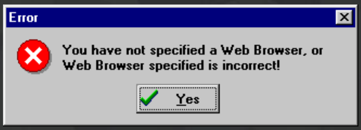
○ Mailto
● CGI

Scrolling Status Bar Message (max length = 200 characters)
WebMania 1.5b with Image Map Wizard is here!!



- H2-8: Aesthetic & minimalist design
  - no irrelevant information in dialogues

Heuristics (cont.)



You have not specified a Web Browser, or Web Browser specified is incorrect!

- H2-9: Help users recognize, diagnose, & recover from errors
  - error messages in plain language
  - precisely indicate the problem
  - constructively suggest a solution

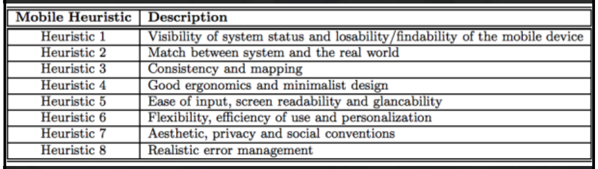
Good Error Messages

- Clearly indicate something has gone wrong
- Be human readable
- Be polite
- Describe the problem
- Explain how to fix it
- Be highly noticeable

Heuristics (cont.)

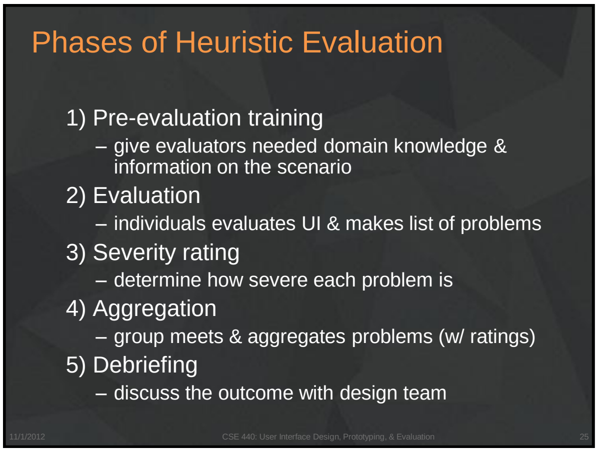
- H2-10: Help and documentation
  - easy to search
  - focused on the user' s task
  - list concrete steps to carry out
  - not too large

Mobile Heuristics



Mobile Heuristic	Description
Heuristic 1	Visibility of system status and losability/findability of the mobile device
Heuristic 2	Match between system and the real world
Heuristic 3	Consistency and mapping
Heuristic 4	Good ergonomics and minimalist design
Heuristic 5	Ease of input, screen readability and glancability
Heuristic 6	Flexibility, efficiency of use and personalization
Heuristic 7	Aesthetic, privacy and social conventions
Heuristic 8	Realistic error management

Enrico Bertini, Silvia Gabrielli, and Stephen Kimani. 2006. Appropriating and assessing heuristics for mobile computing. In Proceedings of the working conference on Advanced visual interfaces (AVI '06). ACM, New York, NY, USA, 119-126. DOI=10.1145/1133265.1133291 http://doi.acm.org/10.1145/1133265.1133291



Phases of Heuristic Evaluation

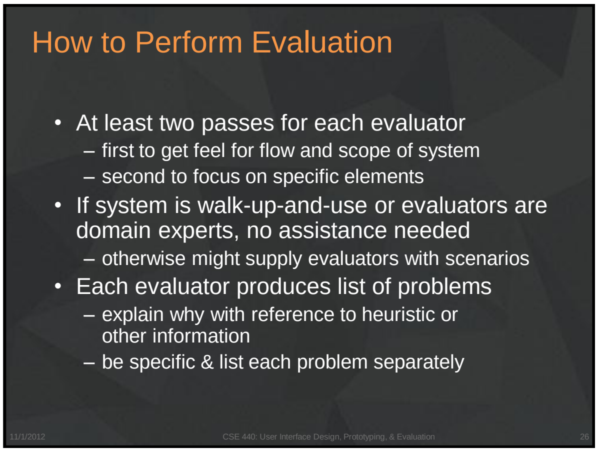
1) Pre-evaluation training
   - give evaluators needed domain knowledge & information on the scenario

2) Evaluation
   - individuals evaluates UI & makes list of problems

3) Severity rating
   - determine how severe each problem is

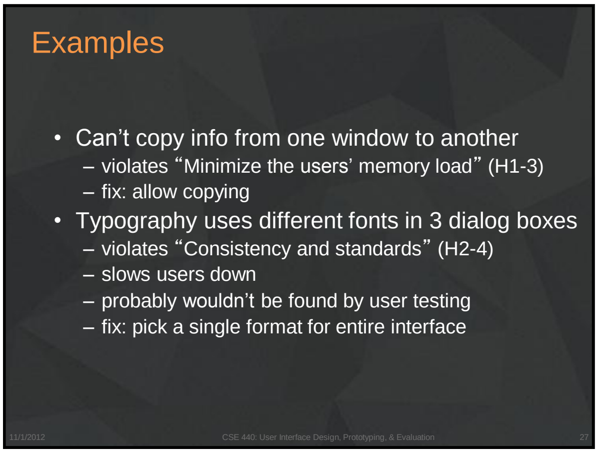
4) Aggregation
   - group meets & aggregates problems (w/ ratings)

5) Debriefing
   - discuss the outcome with design team



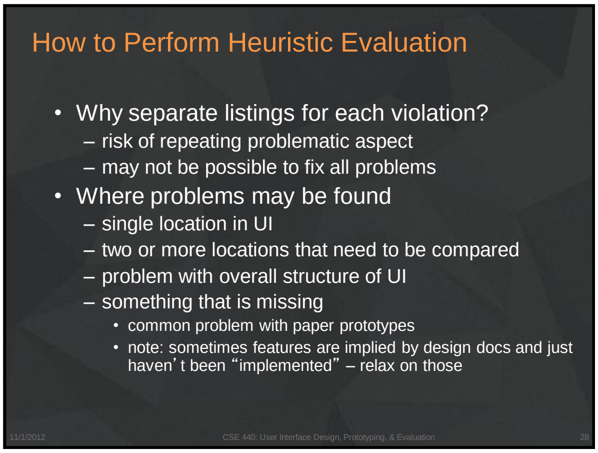
How to Perform Evaluation

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



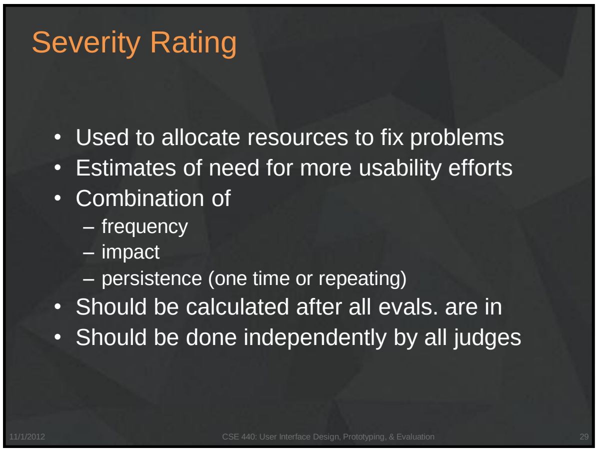
Examples

- Can't copy info from one window to another
  - violates "Minimize the users' memory load" (H1-3)
  - fix: allow copying
- Typography uses different fonts in 3 dialog boxes
  - violates "Consistency and standards" (H2-4)
  - slows users down
  - probably wouldn't be found by user testing
  - fix: pick a single format for entire interface



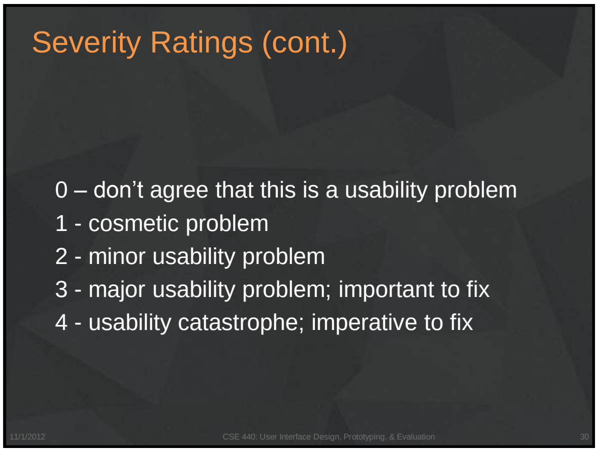
How to Perform Heuristic Evaluation

- Why separate listings for each violation?
  - risk of repeating problematic aspect
  - may not be possible to fix all problems
- Where problems may be found
  - single location in UI
  - two or more locations that need to be compared
  - problem with overall structure of UI
  - something that is missing
    - common problem with paper prototypes
    - note: sometimes features are implied by design docs and just haven' t been "implemented" – relax on those



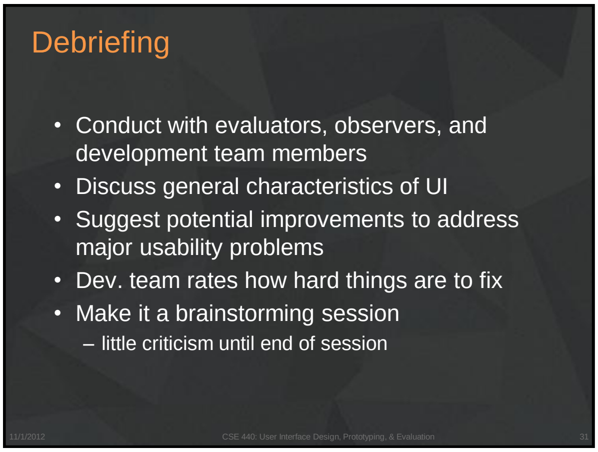
Severity Rating

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



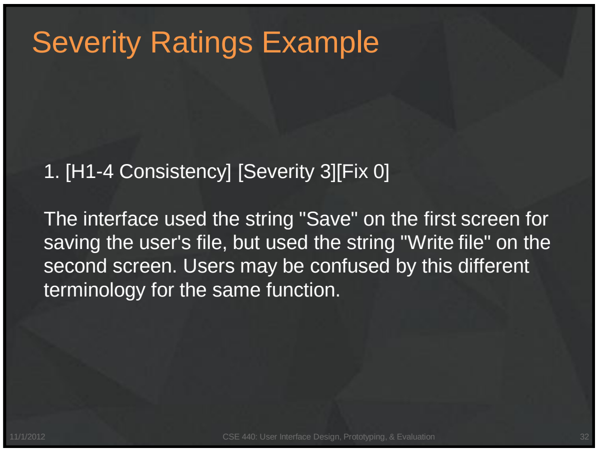
Severity Ratings (cont.)

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



Debriefing

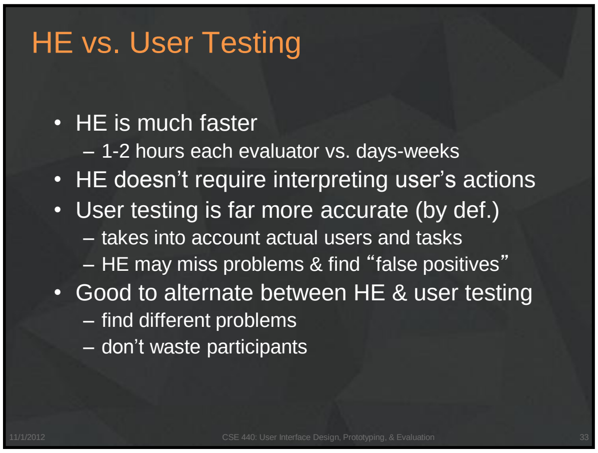
- Conduct with evaluators, observers, and development team members
- Discuss general characteristics of UI
- Suggest potential improvements to address major usability problems
- Dev. team rates how hard things are to fix
- Make it a brainstorming session
  - – little criticism until end of session



Severity Ratings Example

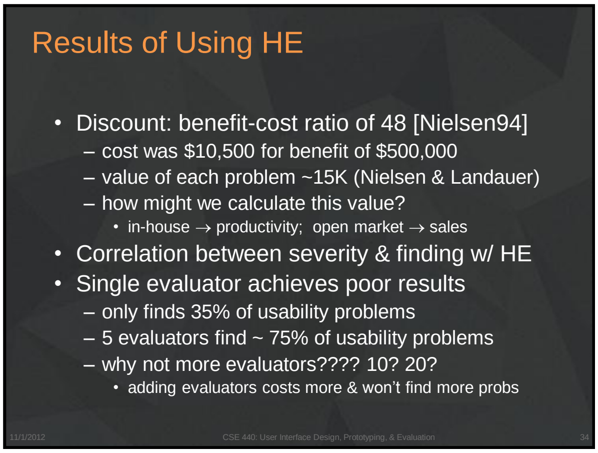
1. [H1-4 Consistency] [Severity 3][Fix 0]

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



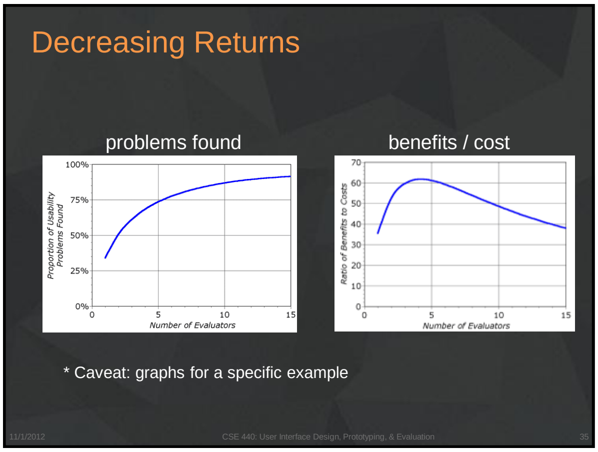
HE vs. User Testing

- HE is much faster
  - – 1-2 hours each evaluator vs. days-weeks
- HE doesn’t require interpreting user’s actions
- User testing is far more accurate (by def.)
  - – takes into account actual users and tasks
  - – HE may miss problems & find “false positives”
- Good to alternate between HE & user testing
  - – find different problems
  - – don’t waste participants



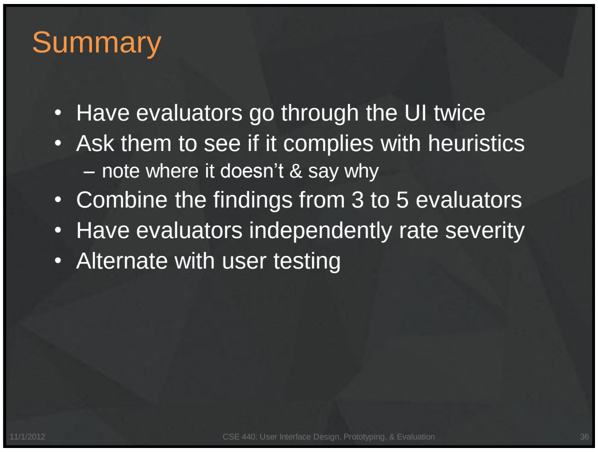
Results of Using HE

- Discount: benefit-cost ratio of 48 [Nielsen94]
  - – cost was $10,500 for benefit of $500,000
  - – value of each problem ~15K (Nielsen & Landauer)
  - – how might we calculate this value?
    - in-house → productivity; open market → sales
- Correlation between severity & finding w/ HE
- Single evaluator achieves poor results
  - – only finds 35% of usability problems
  - – 5 evaluators find ~ 75% of usability problems
  - – why not more evaluators???? 10? 20?
    - adding evaluators costs more & won’t find more probs



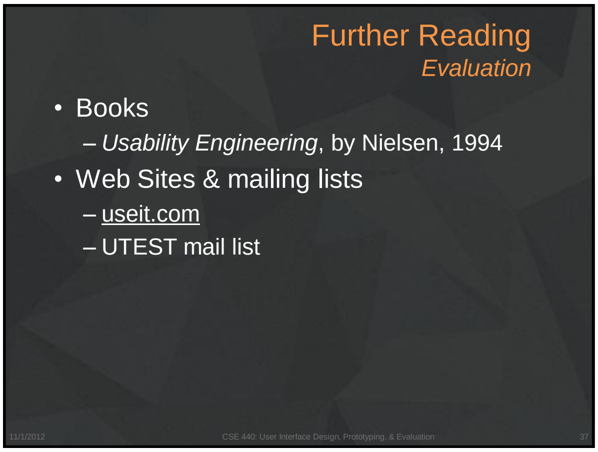
Decreasing Returns

* Caveat: graphs for a specific example



Summary

- Have evaluators go through the UI twice
- Ask them to see if it complies with heuristics
  - – note where it doesn’t & say why
- Combine the findings from 3 to 5 evaluators
- Have evaluators independently rate severity
- Alternate with user testing



Further Reading
Evaluation

- Books
  - *Usability Engineering*, by Nielsen, 1994
- Web Sites & mailing lists
  - useit.com
  - UTEST mail list

11/1/2012 CSE 440: User Interface Design, Prototyping, & Evaluation 37