CS3205: HCI in SW Development

Evaluation (Return to…)

We've had an introduction to evaluation. Now for more details on…

Topics and Readings

- Topics:
  - ”Quick and Dirty” evaluation (?!?!)
  - Using experts for inspections and walkthroughs
- Readings:
  - ID Book, chapter on "Evaluation: Inspections, Analytics, and Models.”
  - 4th edition: Chapter 15 in published book, but no questions on Section 15.3 on Analytics.
  - 3rd edition: Also Chapter 15. No questions on Section 15.3 or Sections 15.4.1 through 15.4.3.
  - See two links at end of main course page.

Reminder: Formative vs. Summative Eval.

- Summative
  - After the something is complete
  - Does it “meet spec”?
- Formative
  - As something is being developed
  - Begin as early as possible
  - For the purpose of affecting the process, the item being developed

High-level Categories of Techniques

- observing users,
- asking users for their opinions,
- asking experts for their opinions,
- testing users' performance
- modeling users' task performance

What You Can Do For Each Type of Evaluation

Method	Controlled settings	Natural settings	Without users
Observing	x	x	
Asking users	x	x	
Asking experts		x	x
Testing	x		
Modeling			x

Quick and dirty

- **‘quick & dirty’ evaluation** describes the common practice in which designers informally get feedback from users or consultants to confirm that their ideas are in-line with users’ needs and are liked.
- Quick & dirty evaluations are done any time.
- The emphasis is on fast input to the design process rather than carefully documented findings.

Quick and Dirty cont'd

- Applies to various other approaches. E.g., could either be:
  - small number of experts (i.e. predictive evaluation)
  - small number of subjects (i.e. usability testing)
- How many subjects? Not as many as you think!
  - See Nielsen's data on why 5 users is probably enough: http://www.useit.com/alertbox/20000319.html
  - 15 users to find 100% of problems. 5 users find 85%.
  - Better three studies with 5 users than one with 15?

Predictive evaluation

- Experts apply their knowledge of typical users, often guided by heuristics, to predict usability problems.
  - **Heuristic evaluation**
  - **Walkthroughs**
- Another approach involves theoretically based models.
  - Predicting time, errors:
  - Example: Fitts' Law formula
- A key feature of predictive evaluation is that users need *not* be present
- Relatively quick & inexpensive

Overview: Evaluation By Experts

- Several approaches
  - Heuristic evaluation
  - Walkthroughs (several flavors)

- In general:
  - Inexpensive and quick compared to asking users
    - "Discount evaluation"
  - Experts may suggest solutions (users probably don't)

Asking experts

- Experts use their knowledge of users & technology to review software usability
- Expert critiques (crits) can be formal or informal reports
- **Heuristic evaluation** is a review guided by a set of heuristics
- **Walkthroughs** involve stepping through a pre-planned scenario noting potential problems

Heuristic evaluation

- Developed Jacob Nielsen in the early 1990s
- Based on heuristics distilled from an empirical analysis of 249 usability problems
- These heuristics have been revised for current technology, e.g., various lists for websites
  - See Box 15.1 in 4th edition of textbook
- Heuristics developed for mobile devices.
  - But for wearables, virtual worlds, etc.?
- Design guidelines form a basis for developing heuristics

Nielsen's "general" heuristics

(Remember these? Also in Chapter 15 of ID Book)

- Visibility of system status
- Match between system and real world
- User control and freedom
- Consistency and standards
- Help users recognize, diagnose, recover from errors
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help and documentation

Heuristics are not Bullet Items

- What you see here in slides are not heuristics usable for evaluation, i.e. not checklists
- For a better example, see: ftp://ftp.cs.uregina.ca/pub/class/305/lab2/example-he.html

1. Visibility of System Status

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

#	Review Checklist	Yes No N/A	Comments
1.1	Does every display begin with a title or header that describes screen contents?	O O O	
1.2	Is there a consistent icon design scheme and stylistic treatment across the system?	O O O	
1.3	Is a single, selected icon clearly visible when surrounded by unselected icons?	O O O	

Discount evaluation

- Heuristic evaluation is referred to as *discount evaluation* when 5 evaluators are used.

- Empirical evidence suggests that on average 5 evaluators identify 75-80% of usability problems.
  - Nielsen (again!). See section 15.2.1 of ID book.

- Compare to "quick and dirty" user studies results (also by Nielsen)
  - Trying to assess same thing: what proportion of errors will this technique find?
  - Each one found "5" is a good number to find a large proportion
    - For the earlier result, 5 users. For this result, it's 5 evaluators.
  - But the proportion of errors that "5" finds is slightly different.

3 stages for doing heuristic evaluation

- Briefing session to tell experts what to do
- Evaluation period of 1-2 hours in which:
  - Each expert works separately
  - Take one pass to get a feel for the product
  - Take a second pass to focus on specific features
- Debriefing session in which experts work together to prioritize problems

Advantages and problems

- Few ethical & practical issues to consider
- Can be difficult & expensive to find experts
- Best experts have knowledge of application domain & users

- Biggest problems
  - important problems may get missed
  - many trivial problems are often identified
- One study has shown:
  - For each true problem, 1.2 false alarms and 0.6 missed problems

Topic: Heuristics for…

- What are a good set of heuristics for a cellphone's UI?
- Status
  - ??
- Navigation
  - ??
- Error prevention
  - ??
- Efficiency
  - ??

Topic: Heuristics for…

- POSSIBLE ANSWERS for What are a good set of heuristics for a cellphone's UI?
- Status
  - Phone should use color or borders to support status.
  - Phone should clearly display battery life / signal strength / time.
  - Phone should display call status (ringing, connected, disconnected)
- Navigation
  - Never be more than 2-3 steps from making call.
  - If move to bottom of screen with many steps, one step back to top.
  - Easy to go back to previous screen / application (undo mistaken selection)
  - See active applications and move to running apps

Topic: Heuristics for…

- POSSIBLE ANSWERS for What are a good set of heuristics for a cellphone's UI?
- Status
  - status of call should be visible (call, connection, roaming, battery)
  - mode (vibrate etc.)
  - unread text messages, voice mails
- Navigation
  - one-button for phonebook numbers
  - consistent navigation button for back, etc.
- Error prevention
  - prevent accidental button presses in pocket, backpack, purse
- Efficiency

Overview: Walkthroughs

- Like heuristic evaluation because
  - Experts are involved
  - Criteria are used to evaluate things
- Different because:
  - Defining characteristic: they “walk through” one or more tasks
  - In addition to experts, may involve designers and/or users

Cognitive Walkthroughs

- Focus on ease of learning
- Designer presents an aspect of the design & usage scenarios
- One of more experts walk through the design prototype with the scenario
- Expert is told the assumptions about user population, context of use, task details
- Experts are guided by 3 questions (on next slide)

- Disadvantages? time-consuming, laborious, narrow focus (maybe that's OK)

The 3 questions

1. Will the correct action be sufficiently evident to the user?
2. Will the user notice that the correct action is available?
3. Will the user associate and interpret the response from the action correctly?

As the experts work through the scenario they note problems

- (See book's link to http://www.userfocus.co.uk/articles/cogwalk.html for 4 questions, similar to these.)

Pluralistic walkthrough

- Variation on the cognitive walkthrough theme
- Performed by a carefully managed team
  - Experts, users, and developers
- The panel begins by working separately
  - Goes through a task scenario, using screens from a prototype (perhaps)
- Then there is managed discussion that leads to agreed decisions
- The approach lends itself well to participatory design
- Disadvantages: larger group to schedule, only look at a few scenarios

Key points

- Expert evaluation: heuristic & walkthroughs
  - Relatively inexpensive because no users
  - Heuristic evaluation relatively easy to learn
  - May miss key problems & identify false ones
  - Walkthroughs: more task focused, more time and cost