EVALUATION AFTER DATA COLLECTION: ANALYZING THE DATA

• Major decision: Accept as is or redesign

  * Decision must be made at a global (metaphor) as well as detailed (individual problem) level

• Compare usability specifications to observed results
  * How close is design to meeting specs?

• If usability specifications are not met, identify interaction design problems and solve in order of cost and effect on usability

  * In most situations, finding usability problems is bad, but in formative evaluation, finding usability problems is good.

• Structured identification of interaction design problems and potential solutions

• Divergent clusters of data might indicate "hidden" user classes
ANALYZING THE DATA

• Cost/importance table

<table>
<thead>
<tr>
<th>Problem</th>
<th>Imp</th>
<th>Solution(s)</th>
<th>Cost</th>
<th>Prio. Ratio</th>
<th>Prio. Rank</th>
<th>Resolutilon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Problem* — something that has an impact on usability; observed as user interacts

*user did not know to select appt before could delete*
ANALYZING THE DATA

- Cost/importance table (continued)

<table>
<thead>
<tr>
<th>Problem</th>
<th>Imp</th>
<th>Solution(s)</th>
<th>Cost</th>
<th>Prio. Ratio</th>
<th>Prio. Rank</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>User didn't know to select appt before deleting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Importance — effect on usability (independent of cost), based on best engineering guess

- Importance = 5: If interaction feature is mission critical or usability problem is "show stopper" (e.g., user cannot complete key task), expected to occur frequently, causes costly errors, or major source of dissatisfaction

- Importance = 3: If user can complete task, but with difficulty (e.g., caused confusion and required extra effort), or problem was source of dissatisfaction

---somewhat subjective; this problem is 3

- Importance = 1: If problem did not impact task performance or dissatisfaction much (e.g., irritant or cosmetic), but is still worth listing
ANALYZING THE DATA

• Adjustment factors for "Importance"

  * Probability of occurrence

    - Over all affected user classes, how often will users encounter this problem?

    - Example: If task cannot be completed (i.e., importance = 5) but usability problem represents situation that will arise rarely and not critical task, downgrade importance to 4 or 3

Or if moderately significant impact (impt. = 3) but occurs frequently, upgrade to 4.
ANALYZING THE DATA

• Cost/importance table (continued)

<table>
<thead>
<tr>
<th>Problem</th>
<th>Imp</th>
<th>Solution(s)</th>
<th>Cost</th>
<th>Prio. Ratio</th>
<th>Prio. Rank</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>User didn’t know to select appt before deleting</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Solution(s) — proposed changes to solve problem

*allow just click in time slot (not require select), then press delete*

- Design principles and guidelines
- Brainstorming
- Studying other similar designs
- Solutions suggested by users and experts
- Typically not a good solution: More training and/or more documentation
ANALYZING THE DATA

• Cost/importance table (continued)

<table>
<thead>
<tr>
<th>Problem</th>
<th>Imp</th>
<th>Solution(s)</th>
<th>Cost</th>
<th>Prio. Ratio</th>
<th>Prio. Rank</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>User didn’t know to select appt before deleting</td>
<td>3</td>
<td>User can click in time slot, then press delete</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Cost — resources (time, money) needed for each proposed solution

4 hours recoding

- Changes to paper prototypes are minimal cost

Thus, cost savings value of paper before computer prototype

- Significance is in computer-based prototypes, versions of real product

* Priority ratio — metric for sorting by priorities:
  (importance rating/cost in person-hours) * 1000

Example: (3/4 hrs) * 1000 = 750

Intuitively, should be high number; hi imp., but low cost

* Priority rank — sorted, descending order of ratios

Fill in after compute all priority ratios

* Resolution — final decision made for addressing each problem, after cost/importance analysis
ANALYZING THE DATA

• Cost/importance analysis

  * Remove all "must fix regardless of cost" problems from table

    * These are EXCEPTIONS: e.g., extreme cost of a "must fix" problem would take it to bottom of list. Exceptions can be dictated by corporate policy, management decision, etc.

  * Of remaining problems, determining which, when fixed, will give biggest improvement to usability for least cost

    * Biggest bang for buck

  * Sort by priority ratio

    * To get hi imp/lo cost first; sort from highest ratio to lowest

  * Fix "must fix" problems first

  * Fix high priority rank (high importance/low cost) problems next

  * Also group by type of problem to find common design solution when possible
DRAWING CONCLUSIONS/REDESIGNING

• Resolution — final decision about each observed problem

  * Compare time available for changes (after deducting time for "must fix" list) to time needed for proposed solutions

  * Three lists

    - Problems that definitely will be fixed, starting with "must fix" list

    - Problems that will be fixed, time permitting

    - Problems tabled

• Implement chosen design decisions

• Realization of benefits of evaluation

• Connection of evaluation back into star life cycle
USABILITY ENGINEERING

• Management of decision-making
• Engineering: achievement of specs, *not* perfection
• Control of iterative process
• Knowing when to stop iterating: when usability specifications are met
USABILITY ENGINEERING

• Goal is cost effectiveness

• If schedule to first release is too tight for thorough formative evaluation in lab, use:
  * Usability inspection
  * Rapid usability analysis
    - Isolate most severe problems, "show stoppers"
    - Leave rest until next release
      * Don’t follow up on problems with lower severity ratings (but note for future)

• Usability engineering tools for classification, reporting, storing and retrieving usability problems, analyses, decisions, etc.

• Videoclip: Analysis of Envision evaluation sessions
  * AFTER VIDEOTAPE: Read thru first part of exercise, have them choose roles, go into new teams, then go on break.
TEAM EXERCISE: FORMATIVE EVALUATION

• Goal:

  * To perform very simple formative evaluation, data collection, and analysis

• Activities:

  PREPARATION

  * Assemble in teams

  * Decide roles for your team members:

    - A prototype executor, to move transparencies and provide feedback (person who knows design best and can "play computer")

    - An evaluation leader, to keep experiment moving, to interact with participants, and to record critical incidents (qualitative data)

    - A user performance timer, to time participants performing tasks and/or count errors (quantitative data)

    - Two participants, to trade to another team
- Anyone else can record critical incidents

* Then, trade the two participants for two from another team, forming new teams. The traded participants are now permanently on their new team.

* Stay with your new team while we hand out and explain forms.

**Do this now.**

* As new teams, excuse both your participants from the room momentarily.

**RUN EXPERIMENT**

* Get your prototype “booted up”.

* Bring first participant into "lab", greet them, and explain evaluation session to them.

* Have first participant use your prototype to perform the benchmark tasks for your objective usability specifications. That is, have participant read first task aloud, then perform that task while
thinking aloud. (Don’t count reading of task in task timing.)

* Next have participant read second task aloud and perform it while thinking aloud.

* Have this participant complete questionnaire, and then give them their "reward".

* Now have the second participant perform the tasks and complete the questionnaire. The first participant should stay and observe.

• **Cautions and hints:**

* Team members must not coach participants as they perform tasks.

* Executor must not anticipate participant actions; especially do not give the correct computer response for a wrong user action! **Executor must only respond to what participants do** (with the prototype), and may not speak, make gestures, etc.

* **You may not change the design on the fly!**
* As executor moves transparencies in response to participant actions, the timer records times and/or counts errors as user performs task. Evaluation leader records critical incidents.

ANALYZE DATA

* The two participants stay with their new team, and the team compiles results to determine whether usability specifications were met.

* Using cost/importance table, list 2 or 3 usability problems from critical incidents.

* Assign an "Importance" rating, 1 through 5, to some of the observed problems, based on severity.

* Propose "Solutions" (without doing all the work of re-design).

* Assign "Cost" values (in person-hours) to each solution. Values based on computer (not paper) prototype changes.

* Compute "Priority ratios".
* Sort by decreasing ratios to determine "Priority rank" of usability problems.

RESOLUTION

* Make some "management" decisions about which changes to make in the next version.

- **Deliverables:**
  - Summary of quantitative results, written in "Observed Results" column, on usability specification table on transparency (for comparison)
  - Cost/importance table, with "Resolutions" column completed in the form on transparency
  - Someone to give brief report on evaluation results

- **Completed by:**
  
  *About 1.5 hours*