Reporting Usability Results
(Creating effective communication)

A Tutorial for User Friendly 2005
December 18, Shanghai

Whitney Quesenbery
Whitney Interactive Design
www.WQusability.com

Agenda

- Defining the problem
- Industry projects on reporting
- Understanding the audience
- Elements of usability reports
- Putting it together
Defining the Problem

Get the project
Do the work
Share the results
Use results to improve product
Why worry about reporting usability results?

Don’t report...tell the story of what you learned

Stories
- Let one person persuade many
- Let the team think creatively
- Act as a springboard for change

<table>
<thead>
<tr>
<th>Method</th>
<th>Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documents</td>
<td>Low</td>
</tr>
<tr>
<td>Charts and diagrams</td>
<td>Limited</td>
</tr>
<tr>
<td>Rational argument</td>
<td>Limited</td>
</tr>
<tr>
<td>Dialogue</td>
<td>Impractical</td>
</tr>
<tr>
<td><strong>Storytelling</strong></td>
<td><strong>High</strong></td>
</tr>
</tbody>
</table>

Adapted from Stephen Denning - www.stevedenning.com
Industry projects on reporting usability

- Defining international standards
  - IUSR Project
  - ANSI standard for reporting summative usability tests
  - New project on reporting formative usability test

- What should be "standardized"?

About the IUSR Project

- Industry Usability Reporting Project (IUSR)
- Started in 1998 and managed at NIST (US National Institute of Standards and Technology)
- Goal: Increase the visibility of software usability
  - Reduce uncontrolled overhead costs of poor usability
  - Encourage software suppliers and consumer organizations to work together
  - Define a process to support decision-making
- The Common Industry Format (CIF), became a US ANSI standard in 2001, and was approved as an ISO standard in 2005
The Common Industry Format (CIF)

- ANSI/INCITS 354-2001
- Codifies best practice for describing a summative test and reporting the statistical results
  - Defines objective usability measures
  - Focuses on reporting usability metrics
  - Provides a description of the test to allow it to be repeated

---

The CIF Template

The CIF is a template for a summative usability test report

- Title page
- Executive summary
- Introduction
  - Full product description
  - Test objectives
- Method
  - Participants
  - Context of the test
  - Tasks given to participants
  - Test facility
  - Equipment used
  - Experimental design
- Usability Metrics
  - Effectiveness (task completion, errors, assists)
  - Efficiency
  - Satisfaction
- Results
  - Data analysis
  - Tables with performance data
- Appendices
  - Questionnaires
  - Participant instructions
  - Task instructions

Join IUSR to get access to the research and full version of the CIF: www.nist.gov/iusr/
IUSR Formative Usability Reporting Project

- **The need**
  - Most usability work is formative, not summative
  - The CIF was being adapted for reporting formative or informal usability tests

- **Goals**
  - Raise the level of visibility of usability
  - Promote best practices
  - Assist practitioners

- **Two workshops**
  - Boston – October 2004
  - UPA 2005 Workshop

---

**Definition of “Formative Usability Testing”**

- A usability evaluation with:
  - Representative users
  - Realistic tasks
  - Some version (prototype, semi-working model, etc.) of the thing being evaluated

- And where the primary purpose of the evaluation is to guide improvement in the design for future iterations
Workshop groups worked on ...

- Reporting context and audiences
  - What are the contextual issues in usability reporting
- Elements of a report
  - What goes into a report...always, or just sometimes
- Reporting metrics
  - When, and how, to report on quantitative metrics
- Guidelines for reporting
  - How do we make decisions about reporting elements, style or audience

Understanding audience and context

- A report is designed like any other interface!
  - Who is the audience?
  - What is our relationship to that audience?
  - How does this affect what and how we report?

- Do we practice user-centered reporting?
Five contexts in which a report is presented

1. Introducing a team or company to usability
   - Need to explain usability
   - Need to avoid jargon
   - Need to establish credibility of the methods

2. Establishing a new “consulting” relationship
   - Meet a new team, and introducing your approach
   - Learn their existing methodology and show how usability fits
   - Need to establish credibility

3. Working with a team where you have an ongoing relationship

4. Reporting to an executive decision maker

5. Coordinating with other usability professionals

In these contexts, the report must
- Teach
- Show credibility and
- Help gain acceptance for the recommendations

1. Introducing a team to usability
   - Need to explain usability
   - Need to avoid jargon
   - Need to establish credibility of the methods

2. Establishing a new consulting relationship
   - Meet a new team, and introducing your approach
   - Learn their existing methodology and show how usability fits
   - Need to establish credibility
### Continuing an ongoing relationship

**3. Working with a team where you have an ongoing relationship**
- You may be part of the team or a “consultant”
- Need to work within expectations and processes you have established
- Can take some “shortcuts” on areas of agreement or where the whole team participated

**5. Coordinating with other usability professionals**
- You may be sharing usability results across projects or over time
- May need more methodology details to satisfy other professionals
- May need more data to allow comparison

In these contexts, the report must

- Communicate efficiently
- Talk professionally to your colleagues

### Reporting to business executives

**4. Reporting to an executive decision maker**
- Keep it short!
- Focus on actions to be taken: recommendations and decisions that need to be made
- Explain how the work was done, but avoid detailed discussions of methodology
- Emphasize connection between business goals and recommendations
- Consider a presentation instead of a report

This is the most difficult context. The report must

- Be brief and to the point
- Teach (at a high level)
- Speak to business needs
Four relationships between author and readers

- **Report Documents Team Process**
  - Usability report author is part of the team
  - The report:
    - Collects findings from the usability test
    - Documents learn process and decisions
    - Acts as a team "memory" and work tool
  - Might:
    - Justifies decisions
    - Explains usability process to others
    - Reports to management on progress or decisions of team

- **Report Feeds Team Process**
  - Usability report author is related to the team
  - The report:
    - Informs about new findings from the usability test
    - Creates recommendations for the team
    - Suggests next steps in the design and process
  - Might:
    - Needs to accommodate different levels of "immersed" in making recommendations
    - Explain usability process used to others
    - Be used to lead or justify decisions made later by the team
    - Report to management on progress or decisions of team

- **Report Informs and Persuades**
  - Usability report author is external and possibly unrelated to the team
  - The report:
    - Informs and persuades
    - Provides an "outside view" to help the team move forward
  - Might:
    - Needs to stand alone
    - Be read and reacted to by people who are hostile to usability input
    - Be used to report to management on progress or decisions of team

- **Report Provides Historical Record**
  - Think, distance, or relationship involves author from reader
  - The report:
    - Provides retrospective view of the usability test approach, method and findings
  - Might:
    - Explain usability process used in the past
    - Presents a high level of detail on the specifics of the test context
    - Contributes to long-term analysis or decisions
    - Be used as part of an overview to a team or management on ongoing usability activities or results

One report, or a collection of deliverables?

<table>
<thead>
<tr>
<th>PRE-TEST</th>
<th>IMMEDIATELY AFTER TEST</th>
<th>1-3 WEEKS POST TEST</th>
<th>LONG TERM</th>
</tr>
</thead>
</table>
| **TEST MATERIAL**
  - Scripts, surveys and other test materials
| **TEST DATA AND INITIAL ANALYSIS NOTES**
  - Notes, log data from the usability session
  - Quick "meme of findings" (perhaps the notes from a team debrief)
  - Immediate "to do" list
| **REPORTING TO THE TEAM AND OTHER STAKEHOLDERS**
  - Team notes or wordlist
  - Formal report
  - Presentation or executive summary (stand-alone or to present report)
  - Video highlights as part of formal report or presentation
  - Next steps or long-term recommendations
| **ARCHIVAL DOCS**
  - Session audio/video
  - Logs and notes
  - Participant details
  - Scripts
  - Surveys
  - Prototypes

In many situations, the data from a usability test lead to multiple reports (or reporting formats) to satisfy different audiences and different contexts. (Gary Kravitz, June 2005)
Elements of usability reports

- What should be included in a usability report
  - What information?
  - How much detail?
  - How to present it?

- Do different audiences need different information in a report?

What’s in a usability report

The “superset” of elements has more emphasis on description and explanation than the CIF template, and includes more detail on results and recommendations.

<table>
<thead>
<tr>
<th>Title page or front matter</th>
<th>Results and recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive summary</td>
<td>Details of recommendations</td>
</tr>
<tr>
<td>Teaching usability</td>
<td>Metrics</td>
</tr>
<tr>
<td>Test background</td>
<td>Quotes, screenshots and video</td>
</tr>
<tr>
<td>Method and methodology</td>
<td>Conclusions</td>
</tr>
<tr>
<td>Overall test environment</td>
<td>Next steps</td>
</tr>
<tr>
<td>Participants</td>
<td>Appendices</td>
</tr>
<tr>
<td>Tasks and scenarios</td>
<td></td>
</tr>
</tbody>
</table>

The list of all elements is in the back of the workbooks, and published in the paper in UPA’s Journal of Usability Studies
Analysis of sample reports

- The IUSR project analyzed 24 reports (some with more than one document)

- There was a lot of variation
  - No elements appeared in every report
  - Some elements were not used in any of the reports

- But most had some form of
  - List of participants
  - Findings or results
  - Recommendations

A summary of this analysis is published in the UPA Journal of Usability Studies

There was a wide range of styles

- Formal documents
- Documents and presentations with screen shots
- Documents with tables of tasks, findings and recommendations
- Spreadsheets or tables with lists of issues
Organized by page, with lots of screen shots and callouts

Organized by scenario and user task

Recommendation Details: Location of Search

Put an entry box for search on the home page, in a visible position

Why?
- Users first scanned the page, looking for a likely link. If they did not find one, they then looked for an "entry box" where they could type.
- Entry boxes have high attraction, but should not be visually hidden in the header.
- A small link to a search will not be noticed easily.
- The best combination we saw was on Merck Praxis: a set of icons, an alphabetical list and the search box were in close proximity.

Presentation format

Classic document format
Front page (or cover)

- Use the cover or first page to identify the test being reported
  - What was tested
  - When and where was the test conducted
  - Who conducted the test
  - Brief description of the participants

This example (from Ginny Redish) is one of the few that showed details of the test so clearly.

Executive summary

- Use the executive summary to provide a 1-2 page overview of the most important findings and recommendations
  - Briefly define the project
  - Summarize overall usability level of product
  - List recommendations at a high level
  - Include positives!
Teaching usability

- Some reports always include general information to teach readers about usability.
- Useful:
  - Introducing a new team to usability
  - For an executive audience
  - But not for an ongoing relationship

Usability priorities

The usability priorities for this report are shown below. These priorities might reflect their goals for the task at hand (or part of it).

1. Effective: I need to read that it got what I was looking for.
2. Easy to learn: I want to be able to use the site without help or feeling like I am lost.
3. Error-tolerant: I need to be confident that the result is correct.
4. Efficient: I am busy and need to use my time well.
5. Engaging: I want an experience that is unique and compelling.

This report included a brief description of the usability priorities we set for an expert review.

Participants

- Include a summary of the participants that shows:
  - How many people you worked with
  - Demographics (relevant to the project)
  - Totals for different characteristics (especially for high numbers)

- Formats
  - Table or list
  - Summary descriptions
  - Profiles or personas

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Job Title</th>
<th>3.0 Experience</th>
<th>4.0 Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jeannette</td>
<td>Trainer</td>
<td>Very experienced</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>Nicole</td>
<td>Training Developer</td>
<td>Experienced</td>
<td>None</td>
</tr>
<tr>
<td>3</td>
<td>Craig</td>
<td>Trainer</td>
<td>Very experienced</td>
<td>None</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Length of time online</th>
<th>Other usage</th>
<th>Child with</th>
<th>Total</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 years</td>
<td>1-20 per week</td>
<td>-</td>
<td>9</td>
<td>Male</td>
</tr>
<tr>
<td>5 years</td>
<td>1x per week</td>
<td>Age: 40</td>
<td>9</td>
<td>Female</td>
</tr>
<tr>
<td>4 years</td>
<td>Daily</td>
<td>Ethnicity: Indian</td>
<td>9</td>
<td>Hispanic American</td>
</tr>
<tr>
<td>3 years</td>
<td>1x per week</td>
<td>Education: PhD</td>
<td>9</td>
<td>College</td>
</tr>
<tr>
<td>2 years</td>
<td>1x per week</td>
<td>Gender: Male</td>
<td>9</td>
<td>Male</td>
</tr>
<tr>
<td>1 year</td>
<td>Daily</td>
<td>Relationship: Sibling</td>
<td>9</td>
<td>Sibling (Siblings)</td>
</tr>
</tbody>
</table>
Participants

These profiles of users for a poker web site, with a brief description of their style of play, prior background and success criteria described general market research.

- **The Tentative Player**
  - This player may be an experienced gamer, used to online games, but new to poker.
  - They have not yet tried online poker.
  - They might have seen poker on TV or in movies.

- **The Casual Player**
  - This player has tried other online gaming sites, including online poker.
  - They find it enjoyable, but it is still an occasional activity, not a primary method of entertainment.

- **The Avid Player**
  - Although an amateur, this user is an experienced and frequent player.
  - They believe the game is too easy, and find it a stimulating challenge.

- **The Embarrassed Player**
  - This player has always admired the game, but feels they are not good enough.
  - They feel they are not good enough and avoid playing.

To choose Full Tilt Poker, he needs:
- Feel the site is a good match to his style of playing
- Feel engaged by the site (e.g., graphics, software)
- To not encounter any usability or trust barriers during download and initial play

These quick personas summarized the user analysis, putting participants into four groups.

<table>
<thead>
<tr>
<th>Magpie (The Collector)</th>
<th>Unconcerned Competitor</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Oh, look...this is interesting.&quot;</td>
<td>&quot;I don't care, I'll play it anyway.&quot;</td>
</tr>
<tr>
<td>Periodically revisits, always looking for more.</td>
<td>Busy, never has time to play.</td>
</tr>
<tr>
<td>&quot;Alright, but not really interested.&quot;</td>
<td>&quot;I'm here for the competition!&quot;</td>
</tr>
<tr>
<td>Snap judgments, cares more about online skills and methodical.</td>
<td>&quot;I don't care about online skills, I want to win.&quot;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Impatient</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;What's up there?&quot;</td>
</tr>
<tr>
<td>Snap judgments, cares more about online skills and methodical.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deeply Engaged</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;I don't care, I'll play it anyway.&quot;</td>
</tr>
<tr>
<td>Periodically revisits, always looking for more.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Magpie (The Collector)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;I love it. I'm going to win big.&quot;</td>
</tr>
<tr>
<td>Periodically revisits, looking for more.</td>
</tr>
<tr>
<td>&quot;What's up there?&quot;</td>
</tr>
<tr>
<td>&quot;I don't care, I'll play it anyway.&quot;</td>
</tr>
<tr>
<td>&quot;I'm here for the competition!&quot;</td>
</tr>
</tbody>
</table>

These quick personas summarized the user analysis, putting participants into four groups.
Prioritizing severity

- Should problems and recommendations be prioritized on a severity scale?
  - Some reports use formal severity scales
  - Some reports list recommendations in a general “priority order”

- If you use a severity scale define the levels in the report.

Organizing results and recommendations

**By page or feature**
- Group by the location in the product where the problem occurred

**Advantages**
- Easy to see all the problems in one page or screen
- Useful when this maps to developer assignments

**Disadvantages**
- Each item is taken in isolation

**By task or scenario**
- Show problems as they relate to a complete task that the user might need to complete

**Advantages**
- User centered – focused on user path
- Not dependent on implementation
- Easier to find common interaction patterns

**Disadvantages**
- Doesn’t map as well to a work list

**By priority or severity**
- Put the critical recommendations at the beginning of the list

**Advantages**
- Focuses on most critical problems
- Works well for overarching design problems

**Disadvantages**
- It can be hard to assign severity accurately
Reporting problems

- This sample reports problems observed in a key area of the interface, but does not include recommendations.

  - User typed the same words again
    - Changed the order
    - Replaced words with phrases
    - Used same words on a different site
    - Or just searched again
  - Refining fields were not understood
    - They looked at them, and made no choices
    - Even users who cared about dates did not use date ranges
    - The most successful were article types... but even these were infrequently used

Matching problems to recommendations

- In this example of a popular format, each problem is matched to a recommendation, and organized by severity.

<table>
<thead>
<tr>
<th>Usability problem</th>
<th>Recommendation</th>
<th>Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Problem described in Full Detail</td>
<td>Recommended usability solution described in full detail here.</td>
<td>...</td>
</tr>
<tr>
<td>2 - Problem described in Full Detail</td>
<td>Recommended usability solution described in full detail here.</td>
<td>...</td>
</tr>
<tr>
<td>3 - Problem described in Full Detail</td>
<td>Recommended usability solution described in full detail here.</td>
<td>...</td>
</tr>
</tbody>
</table>
Recommendations summary and details

- In this template, one section summarizes the major findings, but all findings are listed in detail by scenario.

**Major findings:**

- **Visual recommendations**

  Graphical presentations of recommended solutions can help when there is not a simple list of actions to fix a problem.

  - The following rough concepts are intended only to elicit some validation in how the home page might be presented. They are not design recommendations, but lead the thought:
    - **Concept 1:** a gateway to focus on directions
    - **Concept 2:**

  - There is no simple recommendation to solve the problem. Changes might include:
    - Moving the inset location function to a second screen, so the fields are not part of the home page.
    - Creating a more sophisticated view for the set location/fields, replacing them with the selected location and a button to change locations.

  - In this template, one section summarizes the major findings, but all findings are listed by scenario.

  - Findings should include positives as well as negatives. It is important to let developers know what is working well so that those aspects stay in the product. Also, developers may be more willing to hear about problems when you also let them know about successes.

  - Graphical presentations of recommended solutions can help when there is not a simple list of actions to fix a problem.

  - Each major finding section may also include cross-links to the later pages on specific scenarios that exemplify the major finding.

  - Visual recommendations

  - Graphical presentations of recommended solutions can help when there is not a simple list of actions to fix a problem.

  - Each major finding section may also include cross-links to the later pages on specific scenarios that exemplify the major finding.
Visual recommendations

- This report was organized entirely around wireframes illustrating recommendations for a new way to organize the interface

A. Create a consistent section title (like in cancer types)
B. Start with a basic, short definition of clinical trials to orient users
C. Provide a one-line description of what the search will find
D. Make a good transition to information below the form, with no "false bottoms."
E. Use left menu area for links to section features
F. Provide a minimal link to information for professionals

Reporting metrics

Prioritized Recommendation

Rationale

Performance on Expected Path
Verbal Data
Nonverbal Data
Practitioner Judgment

Record of exact measurements (frequency and impact) by participant and task

Task Level Measurements
- Completion rates
- Completion times
- Satisfaction

Event Level Measurements
- Event descriptions
- Event analyses
Reporting metrics

- Reports that included metrics also used graphs or some visualization to present the data
  - Highlighting in tables
  - Simple Excel bar charts and pie charts
  - More advanced graphs

- If you use graphs, make sure they are readable and communicate well.

Controversies

- Let’s talk about a few controversies, and see what you think:
  - Is a usability report complete without recommendations?
  - How many users do you need to report a problem?
  - Should you use quantitative values when there were just a few users?

- What do you think?
Do you have to have recommendations?

- Is a usability test report complete if you have simply reported on what you observed, or must you recommend changes?

How many users...

- How many users must experience a problem for it to be a valid finding to report?
  
  Are there some problems that don’t need quantitative validation?
  
  Is number of people who encounter an error a good metric?
### Reporting quantity and statistics

- If you only have a few users, should you report quantitative counts? Does it make sense to report that “50% of the users did…” if there were only six participants?

  What information should be reported as statistics?

### Putting it all together

- **How do you “tell the story” for the most impact?**
  - Consider your audience
  - Consider the goals of the report
  - Tell the business story

- **Why not just use a research report format?**
What is the reporting context?

Shared experiences ............................................................ Report read out of context
Report documents ............................................................. Report persuades
Use "team language" ............................................................ Use formal language
Shared project structure ....................................................... Report must stand alone

Case study: building consensus

- Context
  Two vendors (design and usability) working on a new web site
- Problem
  The design agency did not like the idea of being judged
- Challenge
  How to be sure that everyone would accept the results of the test

Our solution:
A team analysis session right after the test (literally – we started at 6pm)

As a team, we:
1. Agreed on what we saw during the test
2. Agreed on what it meant, what was the source of the problem (interaction, terminology, visual...)
3. Brainstormed a general solution (but did not make any final decisions)

Then:
4. We wrote a report that documented this work (along with other minor issues)
5. The designer created changes based on the general solutions we had agreed on
Let’s create an ideal report

- **Content**
  What elements are the most important, and which ones would you leave out?

- **Information Architecture**
  How would you organize the report?

- **Presentation**
  How would you present findings and recommendations?

- **Media**
  How would you deliver the report?

Resources

- **Industry Usability Reporting Project (IUSR)**
  www.nist.gov/iusr/

- **Reporting Formative Usability Test Results (A UPA Workshop Report)**

- **“Towards the Design of Effective Formative Test Reports”**
  Mary Theofanos and Whitney Quesenbery, UPA Journal of Usability Studies, Issue 1, Volume 1, November 2005, pp. 28-46

- **Steve Denning – The Leader’s Guide to Storytelling: Mastering the Art and Discipline of Business Narrative**
  www.stevedenning.com
Whitney Quesenbery is a user interface designer and usability specialist with a passion for clear communication.

She is an expert in developing new concepts for product designs and has produced award winning multimedia products, web sites, and web & software applications.

Whitney is President of UPA - Usability Professionals’ Association and is a leader in the STC Usability and User Experience Community.

Before she was seduced by a little beige computer into the world of usability, Whitney was a theatrical lighting designer on and off Broadway. The lessons and stories from the theatre stay with her in creating user experiences.

Publications

- **Dimensions of Usability**
  in *Content and Complexity*  
  eds. Michael Albers, Beth Mazur.  
  Erlbaum, 2003

- **Personas and Narrative**
  in *The Persona Lifecycle: Keeping People in Mind During Product Design*  
  by John Pruitt & Tamara Adlin  
  Morgan Kaufmann Press,  
  February 2006

More articles and publications on my web site:

http://www.wqusability.com/publications