Usability Testing: Lessons Learned and Looking to the Future

Bill Albert, PhD
Executive Director, Bentley UXC
walbert@bentley.edu
June 11, 2015
About Myself

- Bentley UXC
- Bentley HFID
- UX researcher
- Journal of Usability Studies
My Presentation

• A (brief) history of usability
• Lessons learned about usability testing
• Looking to the future
• Challenges and solutions
A (Brief) History of Usability
1967
Micheal Scriven writes about formative and summative evaluations in the education literature as applied to student learning and assessment. These terms would later be applied to different types of usability evaluations.

1980
Ericson and Simon publish "Verbal Reports as Data" which focuses on using the Think Aloud Method that would later come to dominate usability tests.

1981
Alphonse Chapanis and colleagues publish "Tutorials for the First-Time Computer User," which describes usability as more of a formative than summative activity. They suggested that observing about five to six users using the software will reveal most of the problems in a usability test.

1983
The Psychology of Human Computer Interaction is published by researchers at Carnegie Mellon and Xerox PARC (Stuart Card, Thomas Moran & Allen Newell). This seminal book explains GOMS and Keystroke Level Modeling based on the earlier work of Taylor, Fitts and Gilbreth. The first CHI Conference was held in Boston as part of ACM's SIGCHI subgroup.

1984
Apple introduces the Macintosh during the 1994 Super Bowl, making the case that ease of use sells.

1985
J. Gould and Clayton Lewis publish the influential paper, "Designing for Usability: Key Principles and What Designers Think." They discuss an early and continual focus on users as well as empirical measurement and iterative design.


Smith and Moser publish 957 guidelines for "Designing User Interface Software."

The Human Factor, by Harry Hersh and Dick Rubinstein, both at Digital Equipment Corporation, is published. It is the first book-length description of human-computer interaction.
1995
Jakob Nielsen publishes the first bi-weekly column on usability on useit.com which continues to this day.

The Usability Professionals' Association holds its first annual meeting in Portland Maine.

1998
Usability becomes a standard embedded in ISO 9241 pt 11.

"Web Navigation: Designing the User Experience" is published, one of a new wave of books using the term User Experience.

Rolf Molich conducts the first Comparative Usability Evaluation (CUE) study.

Separately, Jacobsen, Hertzum, and John publish the "The Evaluator Effect in Usability Tests: Problem Detection & Severity Judgments-the first study document the "evaluator effect."

2001
American National Standards Institute (ANSI) develops the Common Industry Format for Usability Test Reports (CIF)

2003
Observing the User Experience is published.

1994
Jeff Rubin publishes The Handbook of Usability Testing

Cost Justifying Usability by Randolph Bias & Deborah Mayhew is published.

Jim Lewis re-examines Virzi's claims in the paper "Sample sizes for usability studies: Additional considerations" and finds general support for the claim that additional users are less likely to reveal new information but the sample size is dependent on the problem occurrence. He also found that problem severity and frequency were independent.

1996
John Brooke publishes the System Usability Scale (SUS) after 10 years of use in industry.

WebEx is founded in California and goes on to develop screen sharing and conferencing software that will be used in moderated remote usability testing.

2000
Don't Make me Think by Steve Krug is published which brings usability testing to the masses using the same Think Aloud method from Ericson and Simon from 20 years earlier.

2002
The first publications about remote usability testing emerge, including the one by Tom Tullis et al on "An Empirical Comparison of Lab and Remote Usability Testing of Web Sites"
2003
Observing the User Experience is published.

2008
Tom Tullis and Bill Albert publish Measuring the User Experience, the first book dedicated to measuring usability, which for the last decade had turned increasingly qualitative.

2012
UPA changes its name to the User Experience Professionals Association (UXPA). Sauro and Lewis publish the first book on the statistical analysis of usability data: Quantifying the User Experience: Practical Statistics for User Research

2002
First publications about remote usability testing emerge, including the one by Lewis et al. on "An Empirical Comparison of Lab and Remote Testing of Web Sites."

2006
Methods of automating heretofore expensive and time consuming usability studies using software and crowdsourcing are published.

2010
Beyond the Usability Lab: Conducting Large-Scale User Experience Studies is published.
Lessons Learned about Usability Testing
Tangible Business Benefits

• Branding – UX is now a key differentiator
• Profits – impact of small changes
• Reduces support costs
• Increases customer loyalty
Best Practice in the Design Process

- Better design
- Improves design efficiency
- Reduce risk of failure
### Key for Context of Product Use During Data Collection

<table>
<thead>
<tr>
<th>Natural use of product</th>
<th>De-contextualized / not using product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scripted (often lab-based) use of product</td>
<td>Combination / hybrid</td>
</tr>
</tbody>
</table>

---

### Landscape of User Research Methods

#### Behavioral
- Eyetracking
- Usability Benchmarking (in lab)
- Usability Lab Studies
- Moderated Remote Usability Studies
- Unmoderated Remote Panel Studies
- Unmoderated UX Studies
- Ethnographic Field Studies
- True Intent Studies

#### Attitudinal
- Participatory Design
- Focus Groups
- Interviews
- Concept Testing
- Diary/Camera Studies
- Customer Feedback
- Desirability Studies
- Card Sorting
- Intercept Surveys
- Email Surveys

---

**Fills a Gap in Research Methods**

Training is Required

- “Thrown into the role”
- Harder than it looks
- Impact on quality of results and organizational perception

Discount Testing

- Online (self-guided) testing
- Creative ways to get participants
- Unmoderated testing (for large sample sizes)
Adaptive Approach

What is the problem and how big is it?

What are people doing?

What are people saying?

What is the problem, why is it a problem and how to fix it?
Product Flexibility

• Works with almost any product or even service
• Tasks are almost always relevant
• Context of use is critical
Sample Size

Probability of Problem Discovery

- Blue line: Probability (p=.48)
- Red line: Probability (p=.25)

For n=3 to n=8,
Lack of Scientific Rigor

CUE-2
The organisations reported 310 different usability problems. Only two problems were reported by six or more organisations, while 232 problems (75%) were uniquely reported, that is, no two teams reported the same problem. Some of the unique findings were classified as serious. Even the tasks used by most or all teams produced very different results - around 70% of the findings for each of these tasks were unique.

CUE-4
The teams reported 340 different usability issues. Only nine of these issues were reported by more than half of the teams, while 205 issues (60%) were uniquely reported, that is, no two teams reported the same issue. Sixty-one of the 205 uniquely reported issues were classified as serious or critical problems.
Power of a Lab Experience

• Seeing failure “live”
• Brings the team together
• Improves decision making
Simple Labs

- Recording/streaming capability
- Online tools and software
- Remote and mobile options
Biases in Testing

- Biases in planning
- Biases in interactions with participants
- Biases in analysis and presentation

http://uxpajournal.org/the-fox-guarding-the-usability-lab/
Quantification Movement

- Requirement in decision making
- Metrics help prioritize
- Measurement is relevant for formative and summative evaluations
Ecological Validity

- Observing use in the real world provides additional insights
- Biggest pain points in the entire experience
- Easier to see unmet needs and barriers to adoption
Personal Tips for Successful Usability Testing

- Find the right participants, but it isn’t a marketing study
- Spend time on the moderators guide
- Create an event for the team
- Don’t confuse behavior and preferences
- Moderate carefully – everything you do makes a difference
- Reports need to be usable – focus on design recommendations
Looking to the Future
UX is a Growth Field

http://blog.generalassemb.ly/the-3-most-promising-careers-of-2020/
http://money.cnn.com/pf/best-jobs/2012/snapshots/43.html
http://www.computerworld.com/article/2496025/enterprise-applications/ux-specialists-are-hot-commodities.html
Holistic Experience

- Usability vs. UX vs. CX
- Emergence of service design

http://kdsketchbook.tumblr.com/
Emotional Engagement

- New software to identify facial expressions
- Highly dynamic
- What is the range of emotions in product use?

http://www.affectiva.com/solutions/affdex-discovery/
Global UX

• Rapid increase in UX awareness globally

• Shortage of best practices and talent

• Challenges with localization
Big UX Data

• Bring together different types of data – power of triangulation

• Opportunity to see hidden trends or patterns in product usage and satisfaction

• Challenge to get access to data, analyze, implement changes, and validate
New Tools

http://www.sticky.ad/
New Interfaces

• Natural gestures
• Issues around learnability, intuitiveness, and efficiency
• What tools do we have to study the usability of new interfaces?
Persuasive Design

- Foundations in consumer psychology and behavioral economics
- Principles such as scarcity, decoys, power of defaults, limits on choice, value of free
- A/B testing is a key tool
Usability of Things

- Context for use is critical – intelligent design
- Competing mental models
- Issues with trust, privacy/security, usefulness, and consistency
Usability in the Wild

- Observe use in natural context
- Drives innovation - barriers to use and unmet needs
- Validation or “ground truth”
Challenges and Solutions
Challenges

• Finding the right talent
• Establishing our place within the organization
• Demonstrating value to the business
• Becoming comfortable with more data
• Expanding our view – broader customer experience
• Ramping up new domains (healthcare)
Solutions

• Educate ourselves
• Initiate strategic projects
• Go deeper and broader
• Share with each other
• Demand to see the data
• Speak the language of business
Obrigado!

Bill Albert, PhD
walbert@bentley.edu
www.bentley.edu/uxc
@BentleyUXC | @UXMetrics