Beyond the Usability Lab

Conducting Large-scale Online User Experience Studies
# Table of Contents

PREFACE .............................................................................................................. xi
ACKNOWLEDGMENTS ......................................................................................... xiii
DEDICATION ........................................................................................................ xv
AUTHOR BIOGRAPHIES ...................................................................................... xvii

CHAPTER 1 INTRODUCTION ............................................................................. 1

1.1 What is an Online Usability Study? ......................................................... 2
1.2 Strengths and Limitations of Online Usability Testing ....................... 5
  1.2.1 Comparing designs ................................................................. 6
  1.2.2 Measuring the user experience ............................................. 6
  1.2.3 Finding the right participants ............................................. 7
  1.2.4 Focusing design improvements ........................................ 7
  1.2.5 Insight into users’ real experience ..................................... 8
  1.2.6 Where are users going (click paths)? ................................ 9
  1.2.7 What users are saying about their experiences ............... 9
  1.2.8 Saving time and money ....................................................... 9
  1.2.9 Limitations of online usability testing .............................. 9
1.3 Combining Online Usability Studies with Other User Research
  Methods ........................................................................................................ 10
  1.3.1 Usability lab (or remote) testing ........................................ 11
  1.3.2 Expert review .................................................................... 12
  1.3.3 Focus groups ...................................................................... 12
  1.3.4 Web traffic analysis .............................................................. 13
1.4 Organization of the Book ........................................................................ 14

CHAPTER 2 PLANNING THE STUDY .............................................................. 17

2.1 Target Users ............................................................................................ 17
2.2 Type of Study .......................................................................................... 19
2.3 Between-Subjects Versus Within-Subjects ........................................... 22
2.4 Metrics .................................................................................................... 24
  2.4.1 Task-based data .................................................................. 24
  2.4.2 End-of-session data .............................................................. 26
2.5 Budget and Timeline ............................................................................... 27
  2.5.1 Budget .................................................................................. 27
  2.5.2 Timeline .............................................................................. 30
2.6 Participant Recruiting ............................................................................. 34
  2.6.1 True intent intercept ............................................................ 34
  2.6.2 Panels .................................................................................. 35
  2.6.3 Direct and targeted recruiting ........................................... 39
CHAPTER 3 DESIGNING THE STUDY ..................................................... 49
  3.1 Introducing the Study .............................................................. 49
    3.1.1 Purpose, sponsor information, motivation, and incentive .................................................. 50
    3.1.2 Time estimate ................................................................ 51
    3.1.3 Technical requirements .............................................. 52
    3.1.4 Legal information and consent .................................... 53
    3.1.5 Instructions .................................................................. 55
  3.2 Screening Questions ............................................................... 56
    3.2.1 Types of screening questions ...................................... 56
    3.2.2 Misrepresentation checks ........................................... 57
    3.2.3 Exit strategy ................................................................ 58
  3.3 Starter Questions ................................................................ 58
    3.3.1 Product, computer, and Web experience .................... 59
    3.3.2 Expectations .............................................................. 60
    3.3.3 Reducing bias later in the study .................................... 61
  3.4 Constructing Tasks ................................................................. 62
    3.4.1 Making the task easy to understand ......................... 62
    3.4.2 Writing tasks with task completion rates in mind ......... 64
    3.4.3 Anticipating various paths to an answer .................... 67
    3.4.4 Multiple-choice answers ........................................... 69
    3.4.5 Including a “none of the above” option ...................... 71
    3.4.6 Including a “don’t know” or “give up” option ............... 72
    3.4.7 Randomizing task order and answer choices ............. 73
    3.4.8 Using a subset of tasks .............................................. 74
    3.4.9 Self-generated and self-selected tasks ....................... 74
    3.4.10 Self-reported task completion .................................... 77
  3.5 Post-Task Questions and Metrics ............................................ 80
    3.5.1 Self-reported data ..................................................... 80
    3.5.2 Open-ended responses .............................................. 82
  3.6 Post-session Questions and Metrics ........................................ 83
    3.6.1 Overall rating scales ................................................ 83
    3.6.2 Overall assessment tools ......................................... 85
    3.6.3 Open-ended questions .............................................. 86
  3.7 Demographic Questions and Wrap-Up .................................... 87
    3.7.1 Demographic questions ............................................ 87
    3.7.2 Wrap-up ................................................................. 88
  3.8 Special Topics .................................................................... 88
    3.8.1 Progress indicators ................................................ 89
3.8.2 Pausing ................................................................. 90
3.8.3 Speed traps .......................................................... 90
3.9 Summary ............................................................. 91

CHAPTER 4 PILOTING AND LAUNCHING THE STUDY ................. 93
4.1 Pilot Data .............................................................. 93
   4.1.1 Technical checks .............................................. 94
   4.1.2 Usability checks .............................................. 97
   4.1.3 Full pilot with data checks ............................... 98
   4.1.4 Preview of results ........................................... 101
4.2 Timing the Launch .................................................. 102
   4.2.1 Finding the right time to launch ....................... 102
   4.2.2 Singular and phased launches ......................... 104
4.3 Monitoring Results .................................................. 105
4.4 Summary ............................................................. 106

CHAPTER 5 DATA PREPARATION ........................................... 107
5.1 Downloading/Exporting Data ..................................... 107
5.2 Data Quality Checks ............................................... 108
5.3 Removing Participants .............................................. 109
   5.3.1 Incomplete data ............................................. 110
   5.3.2 Participants who misrepresent themselves .......... 110
   5.3.3 Mental cheaters ............................................ 110
   5.3.4 Tips on removing participants ......................... 111
5.4 Removing and modifying data for individual tasks ............. 112
   5.4.1 Outliers ...................................................... 112
   5.4.2 Contradictory responses ................................... 113
   5.4.3 Removing a task for all participants ................. 113
   5.4.4 Modifying task success ................................... 114
5.5 Recoding Data and Creating New Variables ..................... 114
   5.5.1 Success data ............................................... 114
   5.5.2 Time variables ............................................. 115
   5.5.3 Self-reported variables ................................... 116
   5.5.4 Clickstream data ......................................... 117
5.6 Summary ............................................................. 119

CHAPTER 6 DATA ANALYSIS AND PRESENTATION ..................... 121
6.1 Task Performance Data ............................................. 122
   6.1.1 Task success ............................................... 122
   6.1.2 Task times ................................................... 131
   6.1.3 Efficiency .................................................... 136
6.2 Self-Reported Data .................................................. 139
   6.2.1 Rating scales ............................................... 139
   6.2.2 Open-ended questions, comments, and other verbatims ................................................. 146
   6.2.3 Overall assessment tools ................................. 152
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.3</td>
<td>Clickstream Data</td>
<td>154</td>
</tr>
<tr>
<td>6.4</td>
<td>Correlations and Combinations</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>6.4.1 Correlations</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td>6.4.2 Combinations (or deriving an overall usability score)</td>
<td>163</td>
</tr>
<tr>
<td>6.5</td>
<td>Segmentation Analysis</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td>6.5.1 Segmenting by participants</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td>6.5.2 Segmenting by tasks</td>
<td>168</td>
</tr>
<tr>
<td>6.6</td>
<td>Identifying Usability Issues and Comparing Designs</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>6.6.1 Identifying usability issues</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>6.6.2 Comparing alternative designs</td>
<td>171</td>
</tr>
<tr>
<td>6.7</td>
<td>Presenting the Results</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>6.7.1 Set the stage appropriately</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>6.7.2 Make the participants real</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>6.7.3 Organize your data logically</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>6.7.4 Tell a story</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>6.7.5 Use pictures</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>6.7.6 Simplify data graphs</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>6.7.7 Show confidence intervals</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>6.7.8 Make details available without boring your audience</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>6.7.9 Make the punch line(s) clear</td>
<td>176</td>
</tr>
<tr>
<td></td>
<td>6.7.10 Clarify the next steps</td>
<td>176</td>
</tr>
<tr>
<td>6.8</td>
<td>Summary</td>
<td>176</td>
</tr>
<tr>
<td>7</td>
<td>CHAPTER 7 BUILDING YOUR ONLINE STUDY USING COMMERCIAL TOOLS</td>
<td>179</td>
</tr>
<tr>
<td></td>
<td>7.1 Loop11</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>7.1.1 Creating a study</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>7.1.2 From the participant's perspective</td>
<td>182</td>
</tr>
<tr>
<td></td>
<td>7.1.3 Data analysis</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>7.1.4 Summary of strengths and limitations</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>7.2 RelevantView</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td>7.2.1 Creating a study</td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>7.2.2 From the participant's perspective</td>
<td>190</td>
</tr>
<tr>
<td></td>
<td>7.2.3 Data analysis</td>
<td>191</td>
</tr>
<tr>
<td></td>
<td>7.2.4 Summary of strengths and limitations</td>
<td>194</td>
</tr>
<tr>
<td></td>
<td>7.3 UserZoom</td>
<td>194</td>
</tr>
<tr>
<td></td>
<td>7.3.1 Creating a study</td>
<td>195</td>
</tr>
<tr>
<td></td>
<td>7.3.2 From the participant's perspective</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>7.3.3 Data analysis</td>
<td>199</td>
</tr>
<tr>
<td></td>
<td>7.3.4 Summary of strengths and limitations</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>7.4 WebEffective</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>7.4.1 Creating a study</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>7.4.2 From the participant's perspective</td>
<td>204</td>
</tr>
</tbody>
</table>
7.4.3 Data analysis ........................................................... 210
7.4.4 Summary of strengths and limitations ................... 211
7.5 Checklist of Questions ....................................................... 212
7.6 Summary .......................................................................... 215

CHAPTER 8 DISCOUNT APPROACHES TO BUILDING AN ONLINE STUDY ...217
8.1 The Basic Approach ........................................................... 217
8.2 Measuring Task Success ................................................... 219
8.3 Ratings for Each Task ........................................................ 222
8.4 Conditional Logic for a Comment or Explanation ................. 223
8.5 Task Timing ...................................................................... 225
8.6 Randomizing Task Order .................................................... 226
8.7 Positioning of Windows ...................................................... 228
8.8 Random Assignment of Participants to Conditions .............. 234
8.9 Pulling it all Together ......................................................... 236
8.10 Summary .......................................................................... 236

CHAPTER 9 CASE STUDIES ................................................................. 239
9.1 Access-Task Surveys: A Low-Cost Method for Evaluating the Efficacy of Software User Interfaces ................................... 240
9.1.1 Background .......................................................... 240
9.1.2 Access Task Survey tool ....................................... 240
9.1.3 Methodology ........................................................ 241
9.1.4 Results ................................................................ 243
9.1.5 Discussion and conclusions .................................. 249
9.2 Using Self-Guided Usability Tests During the Redesign of IBM Lotus Notes ............................................................... 250
9.2.1 Methodology ........................................................ 251
9.2.2 Results ................................................................ 254
9.2.3 Self-guided usability testing: Discussion and conclusions ........................................................................... 254
9.3 Longitudinal Usability and User Engagement Testing for the Complex Web Site Redesign of MTV.com ........................................... 257
9.3.1 Project background .................................................... 258
9.3.2 Why a longitudinal study design ................................ 258
9.3.3 Task structure ........................................................ 259
9.3.4 Data gathering technology and process ................. 259
9.3.5 Respondent recruiting and incentives .................... 259
9.3.6 Lab study and online data gathering methodology verification ........................................................................... 260
9.3.7 Data analysis........................................................... 260
9.3.8 Results and discussion ................................................. 260
9.3.9 Conclusion ........................................................... 264
9.4 An Automated Study of the UCSF Web Site ....................... 266
9.4.1 Methodology ........................................................ 267
Usability labs are a great tool for learning how relatively small numbers of users interact with a design. But we think it’s time to move beyond that. The technology now exists to cost-effectively collect user experience data from hundreds or even thousands of users. Large-scale user experience research brings together a rich set of qualitative and quantitative data about how users interact with a design and their experience. It helps you improve the usability of a design, decide how to prioritize your effort in the design process, discover which of multiple alternatives provides the best user experience, and even whether you should launch your product or go back to the drawing board. We’re confident it will quickly become an indispensable tool in your user research toolkit.

We’re writing this book to share our experiences with large-scale user experience studies, which we also refer to as online studies. Over the last decade we’ve conducted countless studies, some more successful than others. During this time, we’ve figured out what makes for a successful study. We learned the big things, such as how to design a study and analyze the data. We also learned many little things such as which recruiting method makes the most sense when, how to interpret verbatim comments, and how to create usability scorecards based on a variety of metrics. Not only will you avoid some of our mistakes, but you will end up saving time and money, get more out of your studies, and maybe even fewer gray hairs.

This book is not just for seasoned user experience professionals. It’s also meant for those new to the user experience field, or those who simply have an interest in user experience. This technique is valuable for anyone interested in gaining a holistic perspective on the user experience, with an emphasis on obtaining reliable metrics from a large number of users. We wrote the book to be approachable for people with a wide variety of backgrounds and interests. You don’t have to be an expert researcher or web developer to take full advantage of this approach. We will walk you through all the basics you need to know to run an online study. We rely heavily on examples throughout the book, and have a chapter devoted to different case studies so you can see many of the different ways this research has been used.

We’re not expecting you to read this book cover to cover in one sitting, although we welcome that if you so desire! Rather, we wrote this book to be more of a reference guide throughout the course of conducting an online usability study. We’re expecting you to pick up the book at various times during your study, or simply jump around the book as needed. For example, if you simply want to understand what this technique offers, we encourage you to spend some time
on the Introduction and Case Studies. If you’ve never tried to conduct an online test, you may want to focus on the planning, designing, and launching chapters. If you have a very limited budget, you should check out the discount approaches. Above all else, we want this book to be useful, no matter what your background or situation.

We hope you have a chance to visit our website (www.BeyondTheUsabilityLab.com). Throughout the book we will be referring to our website to view or download relevant studies, demos, tools, and calculators. We will do our best to keep the website current, particularly in light of how quickly this technology changes and more research is published. Finally, if you have any questions or suggestions, we would love to hear from you.

bill@BeyondTheUsabilityLab.com
tom@BeyondTheUsabilityLab.com
donna@BeyondTheUsabilityLab.com
We’d like to thank Mary James, David Bevans, and André Cuello from Morgan Kauffmann for all of your helpful suggestions, direction, and hard work in the making of this book. We owe much gratitude to Jim Berling, Tom Brinck, Joe Dumas, Lawrence Fisher, Susan Fowler, and Caroline Jarrett for reviewing the book thoroughly, with great expertise and insight. Thank you to Cheryl Tullis for contributing your graphic design talent to the concept for our cover.

Thank you to our case study contributors: Kirsten Peters, Bob Schumacher, Michelle Cooper, Elizabeth Comstock, Charles Mauro, Allison O’Keefe-Wright, Cianna Timbers, Nate Bolt, Tony Tulathimutte, Alana Pechon, Melanie Baran, Kavita Appachu, Alex Genov, Tyson Stokes, and Heather Fox. Your practical experiences provide value and clarity to this topic.

We’d also like to thank all of our contacts at the vendor services for providing us with complimentary account access throughout this project: Carol Farnsworth and John Quirbach from Keynote Systems, Marshall Harrison and Nazia Khan from RelevantView, and Alfonso de la Nuez and Xavier Mestres from UserZoom.

Last but not least, thank you to everyone else who contributed their thoughts and experiences, and pointed us to valuable resources throughout this process: Ian Cross from Bentley University, Carol Kinas, Yardena Rand, and Fiona Tranquada from Fidelity Investments, and Jeff Miller from Burke Consulting.

**Bill**

First of all, I would like to thank my coauthors, Tom and Donna. Your hard work, dedication, and sense of humor made writing this book a wonderful experience. I will miss our meetings, especially at Mt. Hood. I would also like to acknowledge my sisters, Elizabeth and Laura, and my mother Sara. You bring out the best in me. Lastly, I would like to thank my family. Arjun and Devika, I am constantly inspired by your creativity, curiousity, and playful spirit. To my wife Monika, I can never thank you enough for all of your support. You always believe in me.

**Tom**

I’d like to thank my coauthors, Bill and Donna, but for different reasons. Bill, you’ve been great at keeping us focused on our goals with this book and getting it done. Donna, you’ve taken on some of the most in-depth material in
this book and handled it very well. And you laughed at my jokes even when Bill
didn’t. It’s been great working with both of you. I’d especially like to thank my
wife, Susan, for your help in editing many of the chapters and in fighting with
the style sheets and other oddities in our manuscript to get them consistent. But
most important, I’d like to thank my wife for your understanding and support.
And for not throwing me out of the house when I said I wanted to work on a
second book.

Donna

I’d first like to thank my Mom and Dad who have always unconditionally
supported me. I’m especially grateful of your encouragement throughout this
process, despite my occasional whiny or panicked phone calls. Thank you to
my sister Suzanne, my Grandma, and my other siblings and family members for
your love and inspiration. I’d also like to thank my friends, especially Liz, who
have spent many, many hours of “work parties” with me in cafes, libraries, and
at kitchen tables—I couldn’t have done it without you. Finally, thanks to Bill and
Tom for your hard work, optimism, and tolerance of my sometimes diffuse writ-
ing style and bad jokes. Oh, and thanks to coffee.
Dedication

Bill
To Monika, Arjun, and Devika, and the memory of my father, Lee

Tom
To the memory of my parents, Frank and Virginia

Donna
To Grandma Tess
Bill Albert

Bill Albert is Director of the Design and Usability Center at Bentley University. Prior to joining Bentley, Bill was Director of User Experience at Fidelity Investments, Senior User Interface Researcher at Lycos, and Post-Doctoral Research Scientist at Nissan Cambridge Basic Research. Bill is an Adjunct Professor in Human Factors in Information Design at Bentley University and a frequent instructor at the International Usability Professional’s Association Annual Conference.

Bill has published and presented his research at more than thirty national and international conferences. He recently coauthored (with Tom Tullis) the first ever book on usability metrics, *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*, published by Elsevier/Morgan Kauffman in 2008. He is on the editorial board for the Journal of Usability Studies.

Bill has been awarded prestigious fellowships through the University of California Santa Barbara and the Japanese Government for his research in human factors and spatial cognition. He received his BA and MA degrees from the University of Washington (Geographic Information Systems) and his PhD from Boston University (Spatial Cognition). He completed a post-doc at Nissan Cambridge Basic Research.

Tom Tullis

Tom Tullis is Vice President of Usability and User Insight at Fidelity Investments and Adjunct Professor at Bentley University in the Human Factors in Information Design program. He joined Fidelity in 1993 and was instrumental in the development of the company’s usability department, including a state-of-the-art Usability Lab. Prior to joining Fidelity, he held positions at Canon Information Systems, McDonnell Douglas, Unisys Corporation, and Bell Laboratories. He and Fidelity’s usability team have been featured in a number of publications, including *Newsweek, Business 2.0, Money, The Boston Globe, The Wall Street Journal*, and *The New York Times*.

Tom received a BA from Rice University, an MA in Experimental Psychology from New Mexico State University, and a PhD in Engineering Psychology from Rice University. During his 30 years of experience in human-computer interface studies, he has published more than 50 papers in numerous technical journals and has been an invited speaker at national and international conferences. He also holds eight U.S. patents. He is coauthor (with Bill Albert) of the book *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. 
Donna Tedesco is a Senior Usability Specialist in the Usability and User Insight team at Fidelity Investments. Among other research methodologies, she has conducted numerous online user experience studies over the last several years. Donna received a BS in Engineering Psychology from Tufts University and an MS in Human Factors in Information Design from Bentley University. She has published and presented multiple papers at national and international conferences, including the Usability Professionals’ Association (UPA) and the conference on Computer-Human Interaction (ACM SIGCHI).