

Implementing Usability Testing Of Technical Documents At Any Company And On Any Budget

2010

Meghan Collins
University of Central Florida

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

STARS Citation

Collins, Meghan, "Implementing Usability Testing Of Technical Documents At Any Company And On Any Budget" (2010). *Electronic Theses and Dissertations*. 4372.

<https://stars.library.ucf.edu/etd/4372>

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.

IMPLEMENTING USABILITY TESTING OF TECHNICAL DOCUMENTS
AT ANY COMPANY AND ON ANY BUDGET

by

MEGHAN A. COLLINS
B.A. University of Central Florida, 2004

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Arts in the Department of English
in the College of Arts and Humanities
at the University of Central Florida
Orlando, Florida

Spring Term
2010

Major Professor: Madelyn Flammia

© 2010 Meghan A. Collins

ABSTRACT

In my thesis I discuss the cost effectiveness of usability testing of technical documents and how any size company with any size budget can implement usability testing. Usability is achieved when the people who use products or technical documents can do so quickly and easily to accomplish their own tasks. Usability testing is best defined as the process of studying users to determine a documentation project's effectiveness for its intended audience. Users are tired of dealing with confusing and unintuitive technical documentation that forces them to either call customer service for help on simple issues or throw out the product in favor of one that is more usable or provides better technical documentation. That is why all technical communicators should include usability testing as part of the technical documentation production cycle.

To help technical communicators understand the importance of usability testing, I discuss the cost effectiveness of usability testing and share ways that companies with large budgets and companies with small budgets can begin incorporating usability testing. Then I provide information on all the steps that are necessary for technical communicators to implement usability testing of technical documentation at their company. Options are presented for everything from bare minimum usability testing with a shoe-string budget with pencils, note pads, and only a handful of users to full scale usability testing in large laboratories with the latest equipment and a wide variety of users. The research provides examples from real companies, advice from experienced technical communicators and usability experts, and research demonstrating how many resources are truly required to benefit from usability testing. By showing technical communicators that usability testing is cost effective and that there are many options for implementing usability testing no matter how large or small their budget is, I hope to

empower technical communicators to start including usability testing as part of the documentation production cycle at their companies.

TABLE OF CONTENTS

CHAPTER ONE: INTRODUCTION.....	1
Defining Usability Testing.....	2
Usability Testing and Technical Communication.....	4
Conclusion	6
CHAPTER TWO: LITERATURE REVIEW	7
Introduction.....	7
Planning Usability Tests	8
Conducting Usability Tests.....	15
Managing the Usability Testing Process.....	23
Evaluating Usability Test Results.....	27
Evaluating Usability Testing Methods.....	31
Conclusion	32
CHAPTER THREE: COST EFFECTIVENESS OF USABILITY TESTING	34
Examples for Large Companies or Budgets	38
Examples for Small Companies or Budgets	41
Ways to Convince Your Company to Fund Usability Testing	43
Conclusion	46
CHAPTER FOUR: PLANNING, CONDUCTING, AND MANAGING USABILITY TESTING AND EVALUATING THE RESULTS AND TESTING METHOD	47
Planning Usability Tests	47
Identifying the Users.....	47
Identifying Distinct User Groups.....	48
Finding Test Subjects that Represent the Users.....	51
Determining the Budget.....	53
Choosing a Testing Procedure that Fits the Budget.....	54
Determining the Schedule.....	55
Writing the Test	57
Setting Goals for the Test	57
Setting Measurements for the Test	58
Protecting Against Data Loss.....	60
Conducting Usability Tests.....	61
Conducting Field Tests	62
Benefits	62
Limitations	63
Implementation	63
Conducting Laboratory Tests.....	65
Benefits	65
Limitations	66
Implementation	66
Survey-Based Usability Testing	67
Creating a Usability Team	69
Creating a Full Time Usability Team	69
Creating a Part Time Usability Team	71

Managing the Usability Testing Process.....	71
Evaluating Usability Testing Results.....	73
Evaluating Usability Testing Methods.....	76
Conclusion	77
CHAPTER FIVE: CONCLUSION.....	79
Contribution to the Technical Communication Field	80
Future Research on the Topic	85
Conclusion	88
REFERENCES	90

CHAPTER ONE: INTRODUCTION

When I began the technical communication masters program at the University of Central Florida (UCF) I was introduced to usability testing and was eager to learn more about it. One of the reasons I became drawn to learning more about usability testing is the same reason I became interested in technical communication: there are too many confusing technical documents. Just suggesting that someone “use the manual” or “look it up on the help file” often elicits groans and outright refusal because of bad memories dealing with confusing instructions, unintuitive indexes, indecipherable diagrams, and other breakdowns in communication. Why is it so difficult for so many creators of technical documents to explain coherently to their audience how to perform steps 1-3? Why doesn't the index in the business scheduling program help files have any of the key words that come to mind when dealing with scheduling reoccurring meetings? Why does the drawing of screw A look exactly the same as the drawing of screw B, which is indistinguishable from the drawing of screw C? Why wasn't the warning in a more visible spot in the manual? Why can't technical documents be more user-friendly? When I learned about usability testing in technical documentation, I saw it as one of the solutions to ridding the world of confusing manuals, diagrams, indexes, and other technical documents and became excited to learn more about the process.

Throughout my coursework at UCF my professors encouraged me to include usability testing as part of my document production projects, thus I learned from experience how to perform small scale usability tests, often field tests, for my school projects. My experience with small scale usability testing in my school projects made me curious about how companies perform usability testing on a larger scale, so I chose usability testing as the topic for a literature review I was assigned in one of my classes. As I learned more about usability testing

laboratories filled with expensive equipment and tests conducted by highly paid usability specialists, I started to wonder how technical communicators whose companies have small budgets could begin adding usability testing to their documentation production cycle. I found research that discussed smaller scale usability testing and was encouraged by this research, as well as my own experience with small scale usability tests of my technical documents for school, because it appeared anybody could afford to conduct usability testing and get effective results that would vastly improve their technical documentation. I developed my thesis around this idea that usability testing of technical documentation can be implemented by any size company on any size budget. Before I delve further into my research on usability testing, I want to answer the questions “What is usability testing?” and “How does usability testing fit into technical communication?”

Defining Usability Testing

Before usability testing can be defined, it is necessary to define usability. To find the definition of usability and of usability testing, I consulted the two most commonly referenced books on usability testing for technical communicators, *A Practical Guide to Usability Testing* by Joseph Dumas and Janice Redish, and *Usability Testing and Research* by Carol Barnum. Dumas and Redish state that “usability means that the people who use the product can do so quickly and easily to accomplish their own tasks” (4). Carol Barnum explains that “a product’s usability is determined by the user’s perception of the quality of the product, based on the user’s ease of use, ease of learning and relearning, the product’s intuitiveness for the user, and the user’s appreciation of the usefulness of a product” (6). These definitions of usability remove all mention of the usual ways that companies measure the success of a product or product documentation such as quality assurance and approval from product or documentation experts—

none of which focus on the user and or guarantee usability. Dumas and Redish explain that “[t]esting functionality means making sure that the product works according to the specifications...[u]sability refers to how people work with the product” (4).

The user-centered definition of usability that Barnum and Dumas and Redish offer provides the foundation for the definitions of usability testing that each include in their books. Barnum defines usability testing as “the process of learning from users about a product’s usability by observing them using the product (9).” Other published authors in the field, such as Saul Carliner, Laura Faulkner, and David Wick, accept this definition of usability and refer to it in their articles; in fact, most articles about usability testing in technical communication refer readers to Barnum’s and Dumas and Redish’s usability testing books. Barnum and Dumas and Redish all agree that usability testing has the following characteristics:

1. The primary goal is to improve the usability of a product. For each test, there must be specific goals and concerns that you articulate when planning the test.
2. The participants must represent real users.
3. The participants do real tasks.
4. The team observes and records what participants do and say.
5. The team analyzes the data, diagnoses the problems, and recommends changes to fix these problems. (Dumas and Redish 22)

In other articles on usability testing these same characteristics are described when the author sums up the usability testing process. Hans Van Der Meij describes usability testing as the recording of users’ behavior, verbal reactions, or a combination of the two to identify users’ needs and propensities (195-6). Saul Carliner says:

In a usability test, people who represent intended users are instructed to perform designated tasks under the guidance of the technical communication product...[o]bservations of their behaviors are recorded and used to assess the effectiveness of the technical communication product and to guide efforts to improve it. (91)

In addition, in *Observing the User Experience*, Mike Kuniavsky addresses each of these characteristics in detail throughout his instructions about how to implement usability testing. Some examples from Chapter 2 include defining the audience and their goals, creating tasks that address those goals, getting people who represent the audience, watching them try to perform the tasks, and using the results to address usability issues with the product or documentation (9-15). More details on each of these characteristics are provided throughout the book as Kunviasky explains how to implement each stage of usability testing.

Usability Testing and Technical Communication

Users are tired of dealing with confusing and unintuitive technical documentation that forces them to either call customer service for help on simple issues or throw out the product in favor of one that is more usable or provides better technical documentation. Carol Barnum explains that “changes in the user population and the increased competitiveness of the marketplace” have resulted in a “need for usable products” (2). Usability testing is key to creating usable products and product documentation. Companies such as Boeing Co. have decided that new software purchased by the company must have gone through usability testing because products that have not gone through usability testing have wasted company time and money (Thibodeau 1). When the product is confusing and the technical documentation is difficult to navigate or lacks the information that user needs, users often interrupt co-workers to get help or spend extra time calling customer service. In *Usability Testing and Research*,

Barnum provides an example of how much money an awkward computer program feature can waste. First, she estimates that ten seconds are wasted every time the feature is used. Next, she calculates how much money each second of the user's time is worth and gives an estimate of how frequently the awkward feature is used in a typical year. The result is that the awkward feature causes the sample user \$260.42 a year in lost revenue by wasting the user's time.

(Barnum 23) The loss of time could be even more severe if the product's documentation is also awkward to use or does not include all the information the user needs to use the computer program feature, thus causing the user to lose even more money each year. Time is money and usability problems in products and their technical documentation can quickly eat into a company's bottom line.

Some technical communicators may argue that their documentation does not have usability issues and there is no need for testing, or that including users in the documentation process is a waste of time because users cannot provide valuable feedback, however experts in usability testing disagree. Leo Lentz and Menno de Jong have the opinion that all technical communicators need to incorporate usability testing. Their opinion is based on their experience that "Not one of the many document evaluation studies we have conducted or supervised found the document in question to be free of problems or to contain only minor flaws" (111). Mike Kuniavsky, a founding partner of Adaptive Path, which is a user experience consulting company in San Francisco, states in *Observing the User Experience: A Practitioner's Guide to User Research*, that in Web and product development it was believed that if a product was at all usable then it would be profitable and accepted by users. The prevailing attitude was that asking users their opinion about the usability of the product was irrelevant and likely to damage the brilliance of the product (3). Kuniavsky explains that "[r]ecent history has clearly proved that model

wrong...[n]ow it's necessary to have a product that's actually desired by people, that fulfills their needs, and that they can actually use" (3).

Conclusion

The growing number of articles in the technical communication field about usability testing, as well as the regular occurrence of discussions about usability testing in technical communication organizations such as the Society of Technical Communicators (<http://www.stcsig.org/usability/>) show that more technical communicators are becoming aware of the importance of usability testing, so why aren't all technical communicators implementing usability testing? To answer this question, a variety of research on usability testing is explored Chapter Two of this paper. The research includes usability testing sources from the technical communication field as well as related fields such as computer software and web design. Each stage of usability testing is examined through the research in an attempt to identify all the concerns that technical communicators have about the usability testing process that are preventing them from implementing usability testing at their companies.

CHAPTER TWO: LITERATURE REVIEW

Introduction

Usability testing is an important but oftentimes overlooked element of technical communication projects. Project managers have to juggle demands for time, quality, and budget, and cutting out usability testing is often seen as an easy way to lighten a project budget or to save time. While the upfront savings in budget and time may be appealing, project managers should consider their company's long term budget and time carefully. Usability testing can cut down on the amount of calls from customers with questions and complaints, which can save the company time and money which would be spent hiring customer service representatives and setting up a phone system for customers. Usability tests also improve the quality of the documents and oftentimes of the product can increase the value of the product by making it user friendly. In addition, usability tests on documentation can often lead to spotting bugs with the actual product, which means the product can often be improved through testing its documents.

Despite the importance of usability testing, it is often left out of the documentation process. To see the information that is available to technical communicators and find some clues as to why usability testing is so often neglected in technical communication, I researched the topic and found a large number of articles and textbooks that have been written on usability testing. Most of the articles I found were published in *Technical Communication Quarterly*. I also found some textbooks that specifically discuss usability testing in the technical communication field. In addition, I went outside of these sources to explore articles and books that discuss usability testing of documents for computer science and other related fields. In this literature review I have organized the sources into the following sections: planning usability

tests, conducting usability tests, managing the usability testing process, evaluating usability test results, and evaluating usability testing methods.

Planning Usability Tests

Many project teams question whether it is worth their time to conduct usability tests when they cannot afford an extensive test or do not have access to a desired number of members from the user audience. Geoff Hart, in his article “Ten Technical Communication Myths,” states that one major myth about usability testing in technical communication is that it is too expensive and difficult to do. While it is true that, statistically speaking, the most accurate testing results come from larger and more diverse testing groups, Hart found in his research that even small testing groups can provide important information and an accurate view of the users’ needs. For the technical writer with absolutely no resources for usability testing, Hart encourages having at least one person who can represent the users test the documentation. Hart explains that “[e]ven though one or two users don’t truly represent your audience, many things that these few users find problematic will pose exactly the same problem for the larger audience” (Hart).

Hart also suggests that by introducing the concept of usability testing by first conducting many small, low-budget tests, a project team can pave the way for larger scale tests. Starting with smaller tests gives the project team experience with testing that can prepare them for more complicated testing, and by starting small the project team can prove the value of the tests to supervisors and thus secure larger budgets and resources for testing in future projects. Hart’s myth-busting information provides a great starting point for project teams who have never conducted usability tests before but would like to incorporate them now. It also provides options for project teams who have an inadequate budget for a specific project and cannot afford a more extensive usability testing process.

Now that the importance of any size usability test has been determined, how does one go about planning and writing the test? Carol Barnum's chapter on planning for usability testing, from her book *Usability Testing and Research*, explains how to set goals and measurements for the test. The two main questions Barnum poses to readers are what do you want to learn from the test and how do you want to measure the results you receive (153)? The project team must decide which issues are most prevalent and must create a list of well-defined questions that address these issues. Barnum states "Once the team understands its goal for testing, it can turn its attention to the users' goals" (154). The team's questions about the documentation or product become the basis for constructing the usability test. The questions get revised into a list of measurable goals, for example a goal that the user will perform a task with a maximum of three clicks. According to Barnum, by creating a list of questions and measurable user goals based on the questions, the project team can find many ways to accurately test their document or product and to solve any problems or errors that occur during the testing.

Saul Carliner, in his article "Character-Based, Task-Based, and Results-Based," provides detailed information on a task-based approach to assessing usability that is very similar to the information that Barnum provides in her book. Carliner explains that "For criterion-referenced assessment to be effective, people must begin the design process by defining observable and measurable objectives for the communication product" (Carliner 91). This concept applies directly to usability testing—in order to glean useful information for improving documentation or products, the project team must construct a test that consists of several specific and measurable objectives or tasks for the user to complete that can show the project team where there may be problems in the documentation or product, such as overly complicated instructions for a task or an error resulting from an incomplete line of code.

One simple way to achieve a measurable outcome is to test small sections of the documentation or product in each test rather than testing the entire product in one test. By testing small sections, the project team can assign the user a specific set of tasks to perform that will pinpoint trouble spots in the product. Another way to incorporate Carliner's task-assessment concept into usability testing is to construct unbiased feedback questions for the user to answer. To make the feedback questions unbiased and thus more measurable, the writer must keep the questions neutral by omitting language that could influence the user's answer, such as, "What is wrong with this section of the manual?" Asking the user to point out what is wrong with the manual will force the user to come up with problems even if they did not actually experience any problems with that particular section.

Overall, Carliner's task-based approach to usability testing provides a simple template for any type of usability test and coincides with Barnum's approach to preparing the tests. Without measurable questions and tasks, task-assessment or usability testing will not produce fruitful feedback for the project team.

Brinkman et al.'s findings in their article "Component-Specific Usability Testing" also support smaller scale testing. Brinkman et al. have determined in this article that fewer participants are required in usability tests of different versions of a component due to an increase in test effectiveness when component-specific usability tests are used. The increase in test effectiveness is from the limited scope of the test so any usability problems that are found can be pinpointed to the specific component that is being tested. This makes it easier for the company to fix usability problems because it saves the company time and money that would have been spent searching for the faulty component. This information could be adapted to testing technical documents to help technical communicators save money on usability testing through reducing

the number of test participants and pinpointing the cause of usability issues more accurately by testing smaller sections of the usability document at a time, perhaps one task at a time rather than an entire chapter at a time. Again, these findings compliment Carliner's by showing that testing small sections of products or of technical documentation can be more effective than testing an entire product or technical document at once because problems can be pinpointed to specific components of a product or specific sections of a document.

In order to write measurable questions and tasks for usability tests, technical communicators must carefully consider the product specifications. In Don Reinertsen's article, "Don't Define A Requirement Until You Know How To Test For It," he explains that product specifications are not usually measurable until a test plan is defined (50). As a result, oftentimes products pass usability tests because their specifications were met. However, since the specifications were not measurable, the product's usability has not actually been ensured and problems often still exist. The easiest way to avoid this problem is to only include requirements in the product specifications that are measurable—Reinertsen suggests asking the question "How will we test this" (50)? Reinertsen also suggests discussing the entire testing process, instead of just the requirements, with users when conducting user interviews about the product. Oftentimes the user may define a requirement differently than the developer, which could affect the way the product must be tested. By following these suggestions earlier in the production process, project teams can ensure that measurable requirements will exist for their products and documentation before they even begin usability test planning.

Another consideration when planning a usability test is to figure out which testing method should be used. In Monique Jaspers's article, "A Comparison of Usability Methods for Testing Interactive Health Technologies: Methodological Aspects and Empirical Evidence," she

discusses the advantages and disadvantages of three popular usability testing methods: the heuristic evaluation, the cognitive walkthrough, and the think aloud. She also identifies situations when one method has an advantage over another method and concludes that no one method consistently tops the rest in all situations. Ultimately, Jaspers' advice to usability testers is to use more than one method and to choose methods that compliment each other so that more usability problems can be found and corrected.

While planning the usability test, it is important to realize that a variety of traits among the user audience members can affect the process and results of a usability test. Hall, Jong, and Steehouder in their article "Cultural Differences and Usability Evaluation: Individualistic and Collectivistic Participants Compared," question whether a user's cultural background affects the usability testing process and results. Thus far most studies in usability testing on the effects of a user's background have focused on other variables such as gender, educational levels, and prior knowledge or experience. Hall, Jong, and Steehouder argue that the cultural background of users should be considered a key variable in usability testing conducted with multicultural user groups. Thus, Hall, Jong, and Steehouder created a study by conducting a website usability test on users from Western Europe and Asia/Africa and comparing the process as well as the results. Both groups consisted of male PhD students who spoke English as a second language, and all the students tested a Dutch science website that was written in English. These measures were taken to try to remove gender and education variables from the study while incorporating users of different cultures. Hall, Jong, and Steehouder also used Hofstede's four item Individualism, or IDV, index to test the participants. Hofstede's IDV index tests whether an individual or an entire culture is individualistic or collectivistic in nature.

The results of the two user groups showed some differences, most notably in observations of user behavior during the test. However, it seemed that users from both groups found or experienced many of the same problems with the website, and Hofstede's IDV index did not show any statistical differences in the user results based on their cultures. These results show that while cultural backgrounds of testers may cause some differences in the usability testing process and results, other factors such as gender and education level appear to have a greater affect on usability testing. However Hall, Jong, and Steehouder believe that their study revealed the need for a better understanding of the many factors that may influence the choice of evaluation methods and their results (501). Hall, Jong, and Steehouder distrust their results because they feel that other factors such as social desirability (the desire to adjust to another culture, in this case the Dutch culture that created the website), the selection process, and the limits of Hofstede's IDV index that they used to evaluate the test results of each user group may have obscured the results. Also, they realized that Hofstede's IDV index for evaluating the results may be biased, and a new way to evaluate the test results may be necessary to get more accurate feedback from this study.

Because of the unforeseen variables that may have skewed the results of the study, there is no clear answer for how usability testing is affected by multicultural users; however it is clear that project teams should keep in mind the cultures of their user groups when planning and conducting usability tests in hopes that their documents and products can be made more user friendly. Also, if a user audience consists of multicultural user groups, the project team may need to select users from each distinct cultural group for the usability test in order to get feedback from each culture and thus make the documentation or product user friendly to a wider audience.

Part of the usability test planning process includes planning ways to protect against data loss. While recording observations during usability tests, there are many ways to inadvertently introduce data loss. Spencer's chapter on data and reporting in Chapter 8 of his book, *Computer Usability Testing and Evaluation*, discusses visual observations and manual recording and brings up the difficulty of preventing data loss due to a variety of causes such as interruptions in the observation or recording process; limitations to the brain's ability to process information; and splitting attention between more than one test subject. Testers must plan their data gathering and reporting ahead of time so that they know what they are looking for and know what constitutes an aberration from the norm and avoid data loss as much as possible (17). Essentially, when planning a test the project team should limit data loss by controlling the causes of it when possible. For example, the project team could assign one observer per user so that data loss from splitting attention between test subjects could be prevented. Another example is that the project team could have the test in a laboratory or other controlled environment to prevent interruptions during the test. In addition, the project team could use equipment such as video cameras or voice recorders in order to reduce data loss due to human error. Data loss prevention should be an important element to consider when planning a usability test.

A common concern in the planning process for usability tests is figuring out how long the testing process will take. A multitude of factors can influence the amount of time needed for testing processes, many of which are discussed in Joseph Dumas and Janice Redish's book *A Practical Guide to Usability Testing*. Some examples of factors that can influence the time usability testing takes are the amount of usability planning that has already been done on the project, the complexity of the project, the extent that the product that will be tested, how many users are necessary for the test and how much time each user needs to spend on the test, and how

many other tasks the project team is doing at the same time (101). These factors vary from company to company and from project to project and must be assessed by each project team in order to plan the amount of time to be spent on a usability test. The bad news is that anyone who is new to managing usability tests will have to assess the factors carefully but may under or over estimate the amount of time needed, while more experienced test managers will be able to estimate based on prior tests and experiences.

Conducting Usability Tests

Percival and Johnson describe IBM's method for testing their complex technological products in Percival and Johnson's article "Network Management Software Usability Test Design and Implementation." The tests take place in the company's laboratories where all the necessary equipment is provided for the users. Usually usability tests are performed on a group comprised mostly of novice users. However, at IBM the technology they test is so complex that only the experts can use most of the products that IBM tests, in which case they only include expert users in their usability testing. IBM's thorough understanding of their user audience and products has helped them to focus their usability testing scope in a way that will more accurately help them identify usability problems in their products. Other companies could improve their own testing processes simply by gaining a further understanding of their product or documentation's user audience.

The usability tests at IBM are large and more complex simply because the products that are being tested are also large and complex; several groups run various aspects of each test such as planning, development, and evaluation because none of the groups have the resources to conduct the tests alone. The groups that collaborate on the testing process are comprised of members of various departments such as Information Development; Development; Market

Analysis and Support; Human Factors; and Assurance. These departments, and many others, take active roles in the usability testing planning, implementation, and evaluation. This approach helps to ensure that all the content related to the product is tested such as the product itself, the reference documentation, and the marketing material.

In addition, an Overall Coordinator is assigned to head the test process and to locate suitable users to participate in the test. The Coordinator even makes the arrangements for the users' transportation to the testing lab and handles all the user release forms. Several members of each of the many departments that help with the test volunteer as test monitors to observe and guide users through the test. Percival and Johnson state "There have been no problems in obtaining monitors from product and information development, assurance, and other related areas because it is recognized that the experience gained in working with customers and observing them working with the product is invaluable" (Percival and Johnson). All of the contributing departments benefit by participating in the usability testing process at IBM, which is something for other companies to keep in mind. Perhaps smaller companies with less complex usability tests would also benefit by including more departments in the usability testing process.

Shorter documentation also requires usability tests, as explained by Anita Wright and Jone Rymer in their article, "The Value of Usability Testing in Document Design," which describes the usability testing process they followed to test the IRS Form 2119 (used to report income on the sale of a house). The form was short and confusing to tax practitioners as well as taxpayers. When conducting the usability test, Wright and Rymer tested each group in order to find the usability problems that both user groups experienced with the form.

To get feedback from tax practitioners, Wright and Rymer relied on telephone surveys with ten practitioners. Afterwards, Wright and Rymer compared their results and found that over

seventy percent of the tax practitioners were making mistakes on a line that affects the results of four subsequent questions, which meant that by making one mistake the tax practitioners were giving incorrect value answers for the majority of the form.

To get feedback from the tax payers, Wright and Rymer conducted individual tests with the tax payers. For the test, they requested that tax payers use the think-aloud method so that they could understand the thought process behind each tax payer's actions while filling out the forms. Each tax payer had a scenario that gave them the information they needed to fill out the form and that provided Wright and Rymer with a way of evaluating the test results against the correct results for the given scenario. After the tax payers finished filling out their forms, Wright and Rymer conducted a structured interview with each tax payer, and thus found which lines the tax payers were having trouble understanding.

Wright and Rymer followed testing procedures that enabled them to gather useful information about the usability problems with the Form 2119 and to pinpoint the exact lines and wording that was confusing the user groups. Once Wright and Rymer understood which lines and phrases were confusing to the user groups, they were able to understand why it was confusing and how the users interpreted the confusing lines. For other small documentation or product usability tests, Wright and Rymer's testing approach may also be an effective way of finding trouble spots.

When considering usability tests, Michael Maddox urges project groups to fight the urge to replace usability testing with a simulated user in his article "The Simulated User." Maddox explains that while simulated human users can provide useful information on tactile product tests, for example, test dummies in car crashes, simulated human user interaction has never been successful and probably never will. The main reason Maddox says simulated human user

interaction will never be successful is because when people interact with products and systems, their actions are the result of a combination of training, experience, culture, expectations, environment, and basic perceptual and cognitive abilities.

The complexities of human interactions are beyond what project groups can reasonably predict, therefore usability testing can reveal many usability problems based on the way users actually interact with the products or documentation rather than based on the way the project group may assume users will interact with the product. This disparity between assumed user problems and actual user problems can be especially apparent when the user group consists of more novice users who must learn how to negotiate the product based on limited prior experience or based purely on instructional documentation. Overall, Maddox admits that even with his expertise with usability testing, in every initial formal usability test he has conducted, users have done something that he did not expect (22). Maddox's article stresses the importance of conducting usability tests with human users and the limitations of simulated users in these more cognitive types of tests.

In Barker's book *Writing Software Documentation*, Barker's chapter about conducting usability tests provides information on all aspects of the usability test, but has especially helpful information about usability testing environments. The two types of testing environments are field tests and laboratory tests. Barker explains the strengths and limitations of each environment.

When conducting a field test, the project team sends one or more members out to observe a user testing the product in his or her workspace, often at an office. This type of testing can produce more accurate results because the user is observed in his or her work environment, which means the project team can see how the user will interact with the document or product on

a day-to-day basis. The project team also gets to see the effect of environmental factors on the usability of the product, such as the size of the work space, the types of sounds that may interfere with day to day work, and the quality of the technical resources available to the user. Barker claims that field tests are also less expensive than laboratory tests, mostly because the project team is spared the expense of setting up a facility full of the necessary equipment for the test.

There are limitations to field tests such as data loss or skewed results from interruptions, distractions, inadequate materials for the test, and an overall lack of control over certain test variables. Errors can also occur because the users feel like “someone is looking over their shoulder” and are uncomfortable by this outside presence in their work area (241). This discomfort can cause the user to perform tasks differently than they usually would or can distract the user and result in more errors or overall difficulty during the test.

When conducting laboratory tests, the project team creates their own testing space and provides all the necessary equipment, then arranges for the users to come to the lab to take the usability test. By providing the test space, the project team can exert more control over testing variables, especially equipment issues, interruptions, and other distractions. They can also devise less obtrusive ways to observe the users such as one way mirrors or hidden video cameras. The increased control over testing variables can make laboratory tests more accurate than field tests, but laboratory tests have some limitations.

The main limitation of laboratory tests is that the users may feel uncomfortable being in the new environment, which can distract them and result in more errors when performing tasks or in an increase in the time it takes to perform tasks. Also, laboratory tests are typically more expensive than field tests because of the cost of creating and maintaining the testing facility and equipment. It is important to remember that the cost may be significantly less at a company that

already has an equipped testing facility that is shared with various departments, so for some project teams the cost of a lab test may not be an issue.

Barker's assessment and explanation of the two major environments for usability testing shows that there are pros and cons to both environments. It is up to each project team to decide which limitations they are willing to accommodate and to be aware of the strengths of each environment when planning and conducting their usability tests.

Another important aspect of conducting usability tests is the number of users chosen to represent the document or product's full user audience. David Caulton's article, "Relaxing the Homogeneity Assumption in Usability Testing," refutes Virzi's commonly referenced model, which states that five test subjects are sufficient to find most usability problems. Caulton explains that "the number of subjects required increases with the number of heterogeneous groups" (1). Essentially, for every different user group that is part of the product or document audience, more test subjects are needed in order to more accurately portray the entire audience. For usability tests of documents or products that have a small number of distinct user groups, fewer test subjects can be chosen for testing because the chances are high that most of the testers of a similar group will find the same problems. Caulton does admit that Virzi's model of using five test subjects would function at the simpler level of finding out that usability problems exist, but would not help with pinpointing each instance of a problem or error. The bulk of the article consists of Caulton's step by step analysis of Virzi's model and the creation of Caulton's own more complicated model, which is a mathematical equation with variables for the different elements such as number of homogeneous user groups and number of heterogeneous user groups.

Essentially, when project teams are figuring out how many users are needed for the usability test, Caulton suggests taking into consideration the number of distinct user groups, the size of the usability test (i.e. the number of sections the project team is testing and the number of tasks the project team wants the user to test), the budget, and the percentage of errors the project team wants the users to find during the test.

Another factor when choosing users for the testing process is the users' knowledge of the product or the type of product being tested. Laura Faulkner and David Wick's study (from their article "Cross-User Analysis: Benefits of Skill Level Comparison in Usability Testing") on the results of usability tests with an equal number of novice users and expert users revealed that the novice users find usability issues concerning what Faulkner and Wick call the "learnability" of the product or document, while the expert group finds issues concerning the efficiency or "optimal use" of the product (Faulkner and Wick). Oftentimes project teams neglect to test a product on expert users, which can cause avoidable problems once the product or documentation is completed and released to the public. One grave example that Faulkner and Wick give concerns a chemotherapy machine that killed or severely injured six patients. Doctors who were experts at entering the client's treatment information into the machine were quicker at entering the information into the machine than the machine was of processing it, thus resulting in the death and injuries of the six patients. One doctor actually tested the machine in this manner after an accident with a patient occurred, thus discovering the cause of the usability problem. (Faulkner and Wick) These accidents were avoidable; the machine's reading error would clearly have been discovered through conducting tests that included expert users. Based on this example, and some more theoretical studies that Faulkner and Wick conduct, it is clear that

project teams should test not only novice users but also expert users in order to discover more of the usability problems that exist with the product or documentation.

Faulkner and Wick advise a pre-screening process for users during selection for testing so that an equal number of novice users and expert users can be included in the test. Essentially, novices and experts are two distinct user groups, and as Caulton suggested, users from each distinct user group should be tested in order to find a higher percentage of the usability problems.

Competitive usability testing is discussed in Mike Kuniavsky's book *Observing the User Experience: A Practitioner's Guide to User Research*. The idea behind competitive usability testing is to use the techniques to get the users' perspective on a competition's product so that any usability strengths or weaknesses can be assessed in the competitor's product and that information used to improve your own product. Kuniavsky suggests that the most effective times to conduct this form of usability testing is during the planning phase of the product (the time when the product requirements are established), before a product is redesigned, and after competitors have made substantial changes to their product (420). These are times when new ideas that are gleaned from the research on the competitor's product can most easily be incorporated into one's own product. This process can save a lot of usability testing time and resources and ensure that your product is keeping up with its competition. Kuniavsky advises doing this form of usability testing and research on a regular basis and incorporating the new ideas or information from this type of testing whenever possible. Competitive usability testing appears to be a great way for project groups to stay current and confident about their products and documentation.

Managing the Usability Testing Process

In their article, “Incorporating Usability Testing into the Documentation Process,” Postava-Davignon et al. wrote about the experiences that a growing company had with usability testing. With the company’s growth came a need for more usability testing. Originally each writing team in the technical communication department was responsible for handling the usability testing of the team’s documents; however the increased workload that came with the company’s growth made it difficult for the writing teams to perform the testing. Documentation team members in this company found that they needed a way to better manage the usability testing process; their solution was to create a full time documentation usability team. The team consists of a group of technical communicators with various types of usability testing experience and one usability testing expert. The team helps technical communicators in the department with usability testing, refines testing procedures, and provides a database that tracks results and strategies from various tests and is available to the entire department (Postava-Davignon et al.). Also, the team has created a guide to usability testing for the entire writing department.

While this article’s advice is clearly intended for a larger company with a well staffed publications department, there are ways that a much smaller company with fewer resources could incorporate various elements of this usability testing management concept. Since a smaller company may not have the available staff to commit full time status to usability testing, the company could nominate a group of employees who have more experience with usability testing to act as usability testing mentors to co-workers and project teams. The mentors could help train staff about usability testing through one-on-one work with staff members or by hosting staff meetings to teach usability testing and to discuss recent usability testing experiences and implications. By incorporating a lower budget and lower resource form of a documentation

usability team, any department can slowly work their way towards a more advanced usability testing process, and in the meantime can greatly improve their current usability testing practices.

Better management and implementation of usability testing may soon become a top criterion for companies according to Patrick Thibodeau's article "Users Begin to Demand Software Usability Tests." In this article, Thibodeau discusses how Boeing Co. has decided that new software purchased by the company must have gone through usability tests that conform to the new national standard for assessing usability tests for software products. The new national standard for usability testing is called the Common Industry Format for Usability Test Reports. This standard outlines a format for reporting test conditions and results and gives user companies enough information about a test to replicate it. The idea is that the standard provides a way to objectively evaluate software (1). This new standard for usability tests was organized by the National Institute of Standards and Technology and may force companies to place more importance on testing their products and documentation, especially if other companies follow Boeing's lead in requiring that products they purchase pass these new standards. Boeing has taken this stance because design flaws can cost millions of dollars in lost time and productivity due to the thousands of end users relying on the products. With this new shift towards consumers placing increased importance in the usability testing of products, project teams can expect their companies to start requiring usability testing for all products and documentation, maybe even increased budgets or resources for the tests.

Website merchants have also discovered the importance of usability. Some sites predicted as much as \$6 billion worth of lost money during holiday shopping in 2000 due to difficult website navigation and checkout functions. (Trager 22). Louis Trager's article "Web Site Usability Tests Arrive" discusses the large influx of customers to companies that perform

usability tests on websites to find problems that may be discouraging customers from making their purchases. The usability testing can cost between \$12,000 and \$850,000 annually, depending on the number of tests the company needs (22). Money could be saved, and made, by website merchants by incorporating usability testing during the creation or construction of the website to prepare it for customers in advance. Usability testing would prevent loss of sales due to usability problems with the site. This article reinforces the importance of incorporating and investing in usability testing during the production phases of products and documentation.

Companies are not just worried about the usability of their websites, which mostly affects their customers, but also of their intranets, which are used by their employees. In his article, "Finding Intranet Solutions: Do You Know The Way To San Jose?," Martin White shares information on intranet usability testing that he learned when he attended an intranet conference. White's first observation is the prevalence of workshops and presentations about usability testing. Companies began to realize that their intranets were not providing employees with the necessary information to perform their jobs because of usability issues, mostly stemming from low planning budgets that did not include usability testing. Companies found that rather than complain about usability issues, employees will find a work-around (19). The employees' work-around response can exacerbate the problem because no one reports the usability problems with the intranet, which results in a less effective intranet and overall work day since the problem is never presented to someone who could fix it.

Another concern is that companies do not feel safe sharing their intranets because of unclear legal issues surrounding the protection of company intranets. Since companies do not feel safe sharing their intranets with other companies, other companies do not have real life sample intranets to use as a model for their own intranets. The main reason companies do not

feel safe sharing their intranet designs is because there may be difficulties with legally protecting/copyrighting intranets. In the end, the reluctance of companies to allow other companies to view their intranets results in less sharing of good and poor intranet user designs. Currently companies have to spend time and money conducting their own usability tests instead of gleaning usability advice from other companies.

Because of the usability problems that many companies are finding with their intranets, more attention is being paid to including usability testing in the planning stages of intranet design as well as after the intranet is introduced to employees so that the intranet can become more efficient and reliable for employees and improve company productivity. This attention to the usability of intranets means that many companies will soon have better management of usability testing.

When managing usability testing, it is important to be aware of the possibility of going overboard or introducing bias into the testing process. Darlene Fichter addresses some common mistakes, or “pitfalls,” that beginner usability testing managers make and shares some warning signs so that these behaviors can be identified and avoided. Two of the pitfalls Fichter mentions are ignoring the evidence and over testing. Testing managers often ignore the evidence of a usability issue by explaining it away or explaining away the user’s behavior (Fichter). Fichter advises that whenever it is tempting to explain away test results that project team needs to stop and assess what is really going on. *Over testing*—testing without a clear purpose—is a pitfall because tests should be carefully planned out so that extraneous information is filtered out; project teams should always have a purpose for what and why they are testing.

Fichter's list of possible pitfalls in the usability testing management process would be useful to anyone who manages usability tests so that the time and resources spent on the testing are not wasted and so that test results do not become biased or skewed.

Evaluating Usability Test Results

Hackos's chapter on usability testing in her book, *Managing Your Documentation Projects*, gives helpful advice on interpreting the test results. She points out the fact that oftentimes when a user finds a small problem with the documentation it often is the result of a larger, global problem that the user does not verbalize or often may not recognize. Hackos states that "By correcting the immediate problem, you may neglect the root cause of the problem, which means, in all probability, that the same type of problem will occur again" (542). Thus, it is up to the project team to go beyond the small, quick fix to see if any larger problems exist that would need to be addressed.

The example that Hackos provides is of a user's complaint that she could not find something in the index. The user included the list of terms she had searched for while trying to find this information. The quick fix would have been to include one or all of these terms into the index. However the project manager found that the larger, global problem was that the writing team had only indexed the headings, not the terms within the text. (543) Therefore, for every problem that comes up in the usability test it is imperative that the project team carefully evaluate all probable causes to come up with the best solution rather than settling for the easiest solution.

Another way to find solutions to usability issues that arise during testing is to focus more on users actions during a test instead of their words. Sheri Rosen interviews one company, in her article "Actions Speak Louder than Words," about the company's recent successful overhaul of

their intranet. The interview took place after the company used a large usability test based on observation to reveal the real problems that users were experiencing. Originally the company tried using surveys to find usability problems, but the survey results showed people's opinion about items rather than revealing usability problems and causes of usability problems. Since the surveys were not helping to communicate the problems that users were experiencing with the intranet, the project team decided to assign one to two observers per user to record the users' interaction with the intranet while completing the test. The resulting notes quickly revealed that the intranet was poorly organized and had an inadequate search engine that most users did not even know existed.

Based on these usability issues, the project team was able to better organize the intranet through better indexes and a cleaner homepage with links to all the departmental sites and improve the search engine. An additional observational usability test after the intranet overhaul showed that users were having a much easier time navigating the sites and finding desired information. This company's experience shows that one effective way to get measurable test results is to combine observation in the usability tests as well as user feedback such as surveys.

When evaluating usability test results, it is important to consider the role that use of color may have played in the users' interactions with the product or documentation. Hans Van Der Meij's article "Examining the Relationship Between Quality Writing and Quality Reading" discusses the affect that color can have on users. The way to measure this affect during usability tests is to record how the user reactions may have varied when performing the same task but on sections that displayed different types of colors. If color was not a variable that was planned for during the construction of the usability test, the project team may still be able to analyze some of its affect by noting any unusual variations in user performance during the test and evaluating the

types of color used in each instance to see if they may have played a role in the users' actions. Colors may affect users of different cultural backgrounds in different ways, depending on the types of objects or moods that each culture attaches to different colors, so this is another way that color may affect usability of documents and products and should be taken into consideration when planning a product or documentation for a multicultural user audience.

The affect of color on the testers is just one variable of the usability testing process that technical communicators must evaluate. Jim Flowers provides advice that would be helpful to project teams about how to format the evaluation of test results in his article "Usability Testing in Technology Education." Flowers' article focuses on integrating usability testing training into the classroom. However, many of the techniques that Flowers suggests for the classroom are also applicable outside of it. For test result evaluations, Flowers suggest that a report be written up that contains all of the test results that revealed a problem with the product or documentation, including interview and survey responses along with observations of users during the test. Next, Flowers says that every problem that is brought up in the testing needs a list of probable causes and a solution, thus the report would include the solutions that the project team created for each usability problem. Flowers also says that the report should contain information about "observed usability problems and the design features of the interface, noting what was found to be unintuitive and why" (18). By recording all of this information in the report, the project group is able to evaluate all problems and their causes and therefore increase the usability of the tested product or documentation.

Hornbaek and Frokjaer's studies on the "evaluator effect," which is the observation that usability observers in similar situations often identify very different usability issues, reveal that using a group of evaluators to interpret usability test results can increase the effectiveness of the

test in some ways. They found that matching answers between the different evaluators and grouping together similar findings results in a shorter list of unique problems. The more usability problems that evaluators can agree on, the more likely the problems will be resolved before the final product is shipped. Hornbaek and Frokjaer advise practitioners to use teams of evaluators to analyze usability problems and interpret usability testing results and to group similar answers together to get a more comprehensive view of the usability of the product or document.

Once the usability problems are identified, technical communicators must decide which problems need to be resolved. Lentz and De Jong discuss the problems that professionals in the technical communication field face when judging the severity of usability problems with documentation. There are several shortcuts that experts inadvertently take when guessing how novices will react to a document or a task that are based on the experts' own experience or knowledge. The result is that experts do not always agree on the severity of a usability problem and can overlook or explain away valid problems that users have with a document. Lentz and De Jong explain that when experts are faced with judging the severity of a usability problem, they need to determine whether the problem impairs the readers' understanding of the document (or section if testing smaller portions at a time), and whether this misunderstanding prevents the reader from getting the primary information they need from the document or if it prevents the authors of the document from achieving their primary objective. If so, then the problem is serious and must be addressed in the revisions.

Judging the severity of usability problems is vital because oftentimes technical communicators must juggle a strict budget with concerns about quality and with time restrictions and therefore may not have the resources to address every usability problem that is revealed

during testing. In this case, only severe usability problems would be addressed in the revisions, making it extremely important that technical communicators accurately judge which problems are severe and which problems are not.

Evaluating Usability Testing Methods

In their article “Determining the Effectiveness of the Usability Problem Inspector: A Theory-Based Model and Tool for Finding Usability Problems,” Andre, Hartson, and Williges discuss the difficulty of evaluating usability testing methods and offer a revised inspection tool and method based on Norman’s Stages of Action model. Andre, Hartson, and Williges say that there is a “lack of a sufficiently broad set of cost-effective usability evaluation tools” and usability evaluation methods, which they feel contributes to the plethora of unusable software programs despite the importance most companies’ now place on product usability. One of their examples is the laboratory test, which can be too expensive a usability evaluation method (UEM) for some companies. In order to provide a solution to the budgetary issues of expensive usability tests such as lab tests, Andre, Hartson, and Williges performed studies to see if usability inspection tools (usability tests performed by experts rather than user audience members such as a document or product walkthrough or a heuristic evaluation) could be paired with a lower budget lab test or other usability test to find the same amount of errors and usability problems as a larger budget lab or other usability test. The study revealed that pairing a low budget usability test with a usability inspection tool was as effective as a larger budget usability test in terms of finding usability problems and errors. This information is useful to project teams who have limited budgets for usability testing.

Some project teams may try to find ways to simplify the usability testing process by relying on past test results to apply to current projects. Saul Carliner’s article, “What Do We

Manage? A Survey of the Management Portfolios of Large Technical Communication Groups,” briefly discusses how some technical communicators have attempted to apply the results from one usability test of a document to all the documentation they write rather than conducting a usability test for each separate documentation project. Some authors think that the insights gleaned from testing one documentation project can apply to all documentation projects. These authors create a list of heuristics and principles of user-centered design based on those limited testing results then apply these heuristics and principles to all future projects without even assessing whether these heuristics and principles really work (Carliner).

In his article, Carliner explains that the attempt to create a usability template that works for every project based off of the insights gained from one particular project is faulty because the project team will never find out if their template actually works for the projects they apply it to. Audiences and their needs can change, and new projects can have different errors or usability issues than previous ones did. These differences in audience, audience needs, and projects can create new usability challenges that would go unnoticed if the usability testing process were skipped. Thus, it is important to avoid trying to streamline the process by relying on old test results for different projects; instead, a usability test should be administered for each project in order to account for the inherent differences between different projects and their various user audiences.

Conclusion

A plethora of sources on usability testing exist to educate technical communicators on usability testing, yet there are still companies that do not perform usability tests on their technical documents. I think part of the problem is that, though many of the sources I found discuss the cost effectiveness of usability testing, none of them put cost effectiveness in the forefront, and I

think that the main concern of companies is cost effectiveness. While everyone wants to provide quality service and products to their customers they also want to make a profit, and both time and money are limited resources and must be spent wisely. In my thesis I show readers that investing in usability testing will bring them a healthy return on their investment and I present ideas on ways to plan, conduct, manage, and evaluate usability tests so readers can begin usability testing of technical documentation immediately.

CHAPTER THREE: COST EFFECTIVENESS OF USABILITY TESTING

In this chapter I discuss how usability testing can be cost effective to companies. My hypothesis is that usability testing of technical documents is cost effective for all budgets, can be implemented by companies of all sizes, and can save companies money.

First I discuss why many companies think usability testing is too expensive. Next I discuss three ways that usability testing of technical documents can actually save companies money. Then I share some ways that large companies or companies with large budgets can save money by implementing large scale usability testing of technical documents. Next I share examples of how small companies or companies with small budgets can save money by implementing small scale usability testing of technical documents. Then I explain how to convince your company to fund usability tests of technical documents.

Usability testing is an important but oftentimes overlooked element of technical communication projects. Project managers have to juggle demands for time, quality, and budget, and cutting out usability testing is often seen as an easy way to lighten a project budget and to save time. The sooner the technical documents can be completed the sooner the writers can begin their next project and bring in fresh revenue for the company.

The costs of usability testing include the cost to pay employees for their time spent on the entire testing process, the expenses of setting up a testing laboratory or traveling for field tests, the expense of reimbursing testers for their time or finding volunteers, and the costs involved with making revisions to the technical documents. The entire testing process requires the company to put more money into the project before seeing any profits from sales. The usability testing process can also affect the timelines for other products the company plans to produce and release, again causing a delay before profits are realized. It is easy to see why managers would

be tempted to skip the usability testing process to save time and money in the short term, however usability testing can save money and increase sales in the long term, thus proving to be a very cost effective practice. In fact, Carol Barnum explains that “a common lesson learned the hard way in the software development industry is that it costs 100 times more to make a change in the next release of a product than to make it at the beginning of the project development cycle” (22). There are many expenses that companies incur later on when usability problems remain in the final version of their products. Usability testing can save companies money by helping them cut back on those expenses.

One way that usability testing saves companies money is by reducing customer support calls and emails. A reduction in customer support calls and emails means companies do not need to invest as much money on help phone lines and support email accounts, as well as on employees to run the phone lines and email accounts (Dumas and Redish 16). Since one usability problem could result in numerous customer calls and emails, each problem that is identified and solved during the production phase can prevent the company from paying for the same problem over and over again.

Reducing usability problems for customers means that customers directly benefit from products that have gone through usability testing. In *Usability Testing and Research*, Carol Barnum gives an example of how much money a sample customer can save based on time savings from a program that went through usability testing versus time wasted on a program that did not go through usability testing. In her example, the sample customer saves \$260.42 a year because the customer spends less time on tasks such as troubleshooting the computer program, calling customer service about problems with the computer program, and completing tasks with the program. For this example Barnum used a very conservative estimate of only ten seconds

delay for each time the customer had to stop in the middle of a task due to a usability issue with a computer program the customer used at work. (23) Barnum stresses that when a product is more usable, it results in time and expense savings for customers, which increases company sales through repeat customers and positive reviews of the product (23).

Patrick Thibodeau's article "Users Begin to Demand Software Usability Tests," backs up Barnum's claim that savings for the customer result in better profits for the company. Thibodeau discusses how Boeing Co. has decided that they will only purchase new software from companies that perform usability tests on their products according to the Common Industry Format for Usability Test Reports. The Common Industry Format for Usability Test Reports is a new national standard for assessing usability tests for software products. This standard outlines a format for reporting test conditions and results and gives companies that use the format enough information about a test to replicate it. The idea is that the standard provides a way to objectively evaluate software (1). This new standard for usability tests was organized by the National Institute of Standards and Technology and may force companies to place more importance on testing their products and documentation, especially if other companies follow Boeing's lead in requiring that products they purchase pass these new standards. Boeing has taken this stance because design flaws can cost millions of dollars in lost time and productivity due to the thousands of end users relying on the products. As other clients begin to place as much importance as Boeing does on the usability testing of products they wish to purchase, companies who practice usability testing of their products and documentation will come out on top over those who choose to skip this vital practice.

Usability testing results in better product reviews and sales. If the product does not come with clear instructions, users may have trouble learning how to use the product and will decide

the product is bad because it is hard to use. High-quality instruction manuals or help files can help people use the product and thus increase the usability of the product resulting in good reviews and more sales (Dumas and Redish 14). Usability testing of technical documentation can reveal usability problems with the actual product, which if identified early enough can be fixed before the product is mass produced and shipped, thus directly increasing quality of the product as well as the documentation and increasing sales.

In “Web Site Usability Tests Arrive,” Louis Trager provides an example of how usability testing can increase a company’s customer base and sales. Trager discusses the large influx of website merchant customers to companies that perform usability tests on websites to find problems that may be discouraging customers from making their purchases. Some of these website merchants predicted as much as \$6 billion worth of lost money during holiday shopping in 2000 due to difficult website navigation and checkout functions. (Trager 22). Usability testing of merchants’ websites can cost between \$12,000 and \$850,000 annually, depending on the number of tests the company needs (22). Website merchants can increase their profits by incorporating usability testing during the creation or construction of the website to prepare it for customers in advance. Usability testing would prevent loss of sales due to usability problems with the site. This article reinforces the importance of incorporating and investing in usability testing during the production phases of products and documentation.

Usability testing of internal company systems, such as company intranets and the intranet technical documentation, can also save a company money even though the testing directly benefits the employees rather than the customers. In “Finding Intranet Solutions: Do You Know The Way To San Jose?,” Martin White shares information on intranet usability testing that he learned from attending an intranet conference in 2003. White’s first observation is the

prevalence of workshops and presentations about usability testing. Companies began to realize that their intranets were not providing employees with the necessary information to perform their jobs because of usability issues, mostly stemming from low planning budgets that did not include usability testing. The companies at the conference found that rather than complain about usability issues, employees will find a work-around (19). The employees' work-around response can exacerbate the problem because no one reports the usability problems with the intranet, which results in a less effective intranet and overall work day since the problem is never presented to someone who could fix it. The time wasted by staff due to usability problems with the company's intranet can result in loss of money to the company. Including time and money for usability testing of employee intranets and other systems as well as the accompanying technical documentation before and after the system is set up can save time and money in the long run by making employees more efficient.

There are many ways that companies can capture the savings that are possible from implementing usability testing of technical documentation. In the next section I share some examples of larger scale testing methods that larger companies or companies with larger budgets can implement.

Examples for Large Companies or Budgets

The large scale usability testing methods I introduce include setting up a testing lab, field testing in a wide variety of customer environments, testing a larger group of users, and creating a full time usability testing team.

Companies with large budgets can set up a usability testing lab to create a controlled environment for usability tests. The lab can be used to test the product and the product's technical documents and the budget for the lab can be shared by the product development team

and the technical communication team. Performing usability tests in a lab gives technical communicators control over the environment, the ability to remove themselves from the environment during the test, the ability to test more than one user at a time, and the ability for more executives or managers to observe the test without interrupting the technical communicators' work (Barnum 20). Usability testing in labs can take place at any point during the development of the product and the product's documentation and enables technical communicators to spot complex usability issues early in the developmental stages. Catching problems early means there is more time for technical communicators to correct the problems before the project deadline.

Field testing in a variety of user environments is another large scale usability testing method that can be very useful. Field testing consists of traveling to the users to observe the users interacting with the product in their own environment, typically an office. The benefits include seeing how the user works with the product, seeing how the user's interaction with the product is affected by the user's work environment, seeing whether the user has access to the technical documentation such as manuals and whether the user makes use of the documentation. (Barnum 20). Field testing can help technical communicators and their companies gain a better understanding of their customer base and their needs.

Companies with larger budgets can benefit greatly by finding a wide variety and large number of users to represent the document or product's full user audience during usability testing. David Caulton's article refutes Virzi's commonly-referenced model, which states that five test subjects are sufficient to find most usability problems. Caulton explains that "the number of subjects required increases with the number of heterogeneous groups" (1). Essentially, for every different user group that is part of the product or document audience, more

test subjects are needed in order to more accurately portray the entire audience. For usability tests of documents or products that have a small number of distinct user groups, fewer test subjects can be chosen because the chances are high that most of the testers of a similar group will find the same problems. Caulton does admit that Virzi's model of using five test subjects would function at the simpler level of finding out that usability problems exist, but would not help with pinpointing each instance of a problem or error. When project teams are figuring out how many users are needed for the usability test, Caulton suggests taking into consideration the number of distinct user groups, the size of the usability test (i.e. the number of sections of the document that the project team is testing and the number of tasks the project team wants the user to test), the budget, and the percentage of errors the project team wants the users to find during the test. By following Caulton's advice and testing for each distinct user group, the product and its documentation will be usable for a greater number of customers. Making a product and its documentation usable to a greater number of customers can result in higher sales and better reviews because more customers can successfully use the product.

Another way companies with large budgets can benefit from usability testing is to create a full-time usability testing team. Postava-Davignon et al. wrote about a company that decided their usability testing needs had grown and had become overwhelming for individual writing teams to handle, especially teams with writers whom had little experience with usability testing. Documentation team members in this company found that they needed a way to better manage the usability testing process; their solution was to create a full time documentation usability team. The team consists of a group of technical communicators with various types of usability testing experience and one usability testing expert. The team helps technical communicators in the department with usability testing, refines testing procedures, and provides a database that

tracks results and strategies from various tests and is available to the entire department (Postava-Davignon et al.). Also, the team has created a guide to usability testing for the entire writing department.

While these techniques are on a larger scale and cost more to implement, there are ways for companies with smaller budgets to adapt those techniques to a smaller scale and still benefit from the usability testing process. In the next section I share ways to scale down the large budget techniques to fit into smaller usability testing budgets.

Examples for Small Companies or Budgets

The small scale usability testing methods I introduce include setting up a temporary testing space, field testing in a few local customer environments, testing a small group of users, and establishing usability testing mentors within the company.

Testing without a lab can be accomplished by setting up a temporary testing space in any room that can fit the equipment needed for the user and seats for the observers and test administrator (Barnum 18). Setting up a video camera to capture the user's actions is optional and can be cut if there are budget constraints. According to Carol Barnum, "At its simplest level, usability testing without a lab requires nothing more than a stopwatch, pencils, some note cards, and some sticky notes" (18).

Postava-Davignon et al. described how a large company set up a full time usability testing team. While it may not make sense for a smaller company or a company with a smaller budget to set up a full time usability team, there are ways that a much smaller company with fewer resources could incorporate various elements of a full time usability team. Since a smaller company may not have the available staff to commit full time to usability testing, the company could nominate a group of employees who have more experience with usability testing to act as

usability testing mentors to co-workers and project teams. The mentors could help train staff about usability testing through one-on-one work with staff members or by hosting staff meetings to teach usability testing and to discuss recent usability testing experiences and implications. By incorporating a lower-budget documentation usability team, any department can slowly work their way towards a more advanced usability testing process, and in the meantime can greatly improve their current usability testing practices.

Testing with a smaller sample of users is another option for companies with small budgets. While David Caulton argues that usability testing is most accurate when sample users from each distinct user group are tested, he does admit that Virzi's model of using five test subjects would function at the simpler level of finding out that usability problems exist, but would not help with pinpointing each instance of a problem or error. However, others in the technical communication field think that Virzi's model is effective for finding the majority of usability problems in a product or document. Jakob Nielsen and Tom Landauer, who determined in their research that "the maximum benefit-cost ratio is achieved when using between three and five subjects," were the first to suggest that such a small number of testers could result in accurate usability testing (Nielsen 251). After Nielsen and Landauer's research was published, Robert Virzi performed his own research and found that "between four and five users 80% of the usability problems are detected and that additional subjects are less and less likely to reveal new information," (291) A few years later James Lewis performed his own study with similar results. Therefore, companies with small budgets can reduce usability testing costs by only gathering five testers and know that their product and documentation will still benefit greatly from the testing.

For companies or technical communicators who cannot gather five testers, Hart's article on the ten common myths in technical communication provides some encouraging news. One of the myths in Hart's article is about usability testing in technical communication. This myth states that it is too expensive and difficult to perform usability testing. While it is true that, statistically speaking, the most accurate testing results come from larger and more diverse testing groups, Hart found in his research that even small testing groups can provide important information and an accurate view of the users' needs. For the technical communicator with absolutely no resources for usability testing, Hart encourages having at least one person who can represent the users test the documentation because many things that one person finds problematic would prove problematic to other users, too.

The large and small scale usability testing techniques can only be implemented if technical communicators have permission from their companies to conduct usability tests. For those technical communicators who must convince their companies to fund usability testing, I have provided some ideas for introducing usability testing to skeptical managers in the next section.

Ways to Convince Your Company to Fund Usability Testing

Technical communicators whose companies do not implement usability testing of technical documentation and who wish to change that will need to convince their companies that usability testing is worth the time and money it requires. Starting with small, low-budget tests, starting with less equipment, sharing testing lab space with other departments, and calculating the savings from usability testing are all ways that technical communicators can introduce the usability testing process to their companies and begin to calculate the savings from implementing usability testing.

Geoff Hart suggests that by introducing the concept of usability testing by first conducting many small, low-budget tests, a project team can pave the way for larger scale tests. Starting with smaller tests gives the project team experience with testing that can prepare them for more complicated testing, and by starting small the project team can prove the value of the tests to supervisors and thus secure larger budgets and resources for testing in future projects. Small, low-budget usability tests are a great starting point for project teams who have never conducted usability tests before but would like to incorporate them now. Experience with small, low-budget testing is useful even for employees in a writing department that usually conducts larger scale usability tests because the department may be forced to cut a project team's budget for a specific project.

Starting with less equipment for usability tests is another way to increase the chances of convincing companies that usability testing is worth the cost. Some items that are useful but not required for usability testing are: a laboratory with one-way mirror, data-logging software, videotape, and a formal test report (Dumas and Redish 25). As previously mentioned, Carol Barnum says "At its simplest level, usability testing without a lab requires nothing more than a stopwatch, pencils, some note cards, and some sticky notes" (18). Reducing the equipment used in the tests can help technical communicators keep the budget for the tests small and can increase their chances of convincing their companies to approve the tests.

At companies that already implement usability testing of products, technical communicators can inquire about sharing existing usability testing labs or about including the technical documentation in the product tests. As Carol Barnum explains, it can be very expensive to setup a testing lab: "[a] fully developed, state-of-the-art usability testing lab costs \$75,000 to \$150,000, depending on the specific needs of the company for testing products" (22).

Using an existing testing lab saves the company the expense of setting up a new one for technical communicators. Also, as Dumas and Redish share “we learned very quickly that in any usability test, even if you are focusing on the documentation, you learn a tremendous amount about the software (and vice-versa)” (29). Since testing the technical documentation also involves testing the product itself and reveals usability issues with both documentation and product, sharing product testing labs or including technical documentation in product tests can provide many benefits and help technical communicators introduce usability testing of documents.

As technical communicators introduce usability testing, they must measure the benefits of usability testing to convince management to continue approving usability testing and to increase the budget available for testing. Dumas and Redish provide one way of calculating the savings resulting from usability testing. Their approach is to “conduct iterative usability tests with the same tasks” (19). The following are the steps in their approach:

1. Conduct an initial usability test, recording times for a set of tasks
2. Change the product based on the results of the test
3. Conduct a second usability test, recording times for the same tasks
4. Subtract the times for the tasks in the second test from the times for the first test (The difference is the time saved because of improvements that are due to the usability engineering approach.)
5. Multiply the time saved by the number of people who use or will use the product
6. Convert the time saved to dollars saved in terms of the users’ average salaries (19).

As Dumas and Redish point out, this approach is very conservative because it does not include savings from reducing customer support calls and emails or making changes to the product after it has been released (19).

Another way to measure the benefits of usability testing is to compare the number of calls to the help desk or technical support department after the release of a new product that was tested for usability to the number of calls to the help desk after the release of a previous product that was not tested for usability (Barnum 50). A drop in the number of calls can provide proof to companies of the benefit of usability testing and help technical communicators secure budget and time for usability testing of technical documents.

Conclusion

In this chapter I discussed three ways that usability testing is cost effective to companies. I also gave examples of ways that large companies or companies with large budgets can implement usability testing, and ways that small companies or companies with small budgets can implement usability testing. Then I shared four ways that technical communicators can convince their companies to implement usability testing of technical documents.

CHAPTER FOUR: PLANNING, CONDUCTING, AND MANAGING USABILITY TESTING AND EVALUATING THE RESULTS AND TESTING METHOD

In this chapter I discuss in further detail ways for large companies or companies with large budgets, and small companies or companies with small budgets to plan and conduct usability tests. I include details on how to determine a budget, create a testing procedure that fits the budget, create a usability testing team, and write the usability test. I also discuss the benefits and weaknesses of laboratory tests, field tests, and surveys, and ways to manage the usability testing process. Throughout the chapter I include ways that companies with any size budget can adapt the information to fit their needs.

Planning Usability Tests

When planning the usability test, technical communicators must begin by identifying the users of the product or documentation they are testing. Identifying the users includes the tasks of identifying distinct user groups, finding and recruiting test subjects that represent users, and determining appropriate compensation for users who take the test. Next, technical communicators must determine the budget for the usability test and choose a testing method that fits the budget. After determining the budget, technical communicators must determine the schedule. Lastly, technical communicators will write the test by setting goals and measurements for the test while considering ways to protect against data-loss.

Identifying the Users

The first step in planning usability testing is to identify the user. For technical communicators the user is the customer who will be using the product that the technical documents support. For example, a bank decides to buy a new computer program that the bank

tellers will use to process transactions. The program will be installed by the IT department. In this scenario, the IT personnel who will install a banking computer program and the bank tellers who will use the program are all users who technical communicators will need to consider while writing the technical documentation for the program. Therefore, it would be helpful during the usability testing to have representatives of both the IT personnel group and the bank teller group and to test each group on the sections of the document that apply to the tasks they will carry out.

Typically, usability tests are performed on a group comprised mostly of novice users. However, at IBM the technology they produce is so complex that usability tests are only performed on a group of expert users since only experts can use most of the products. IBM's thorough understanding of their users and products has helped them to focus their usability testing scope and more accurately identify usability problems in their products (Percival and Johnson). Other companies could improve their own testing processes simply by gaining a further understanding of their product or documentation's users.

Identifying Distinct User Groups

The optimal way to find people for the usability test who represent the greatest number of the product and documentation's users is to follow David Caulton's advice. David Caulton advises technical communicators to find users for the test who represent each distinct user group (1). In the banking software example from above, two distinct user groups were identified based on the different type of tasks they will perform using the same computer program and documentation. There are many other traits used to identify distinct user groups.

Thus far most studies in usability testing on the effects of a user's background have focused on other variables such as gender, educational levels, and prior knowledge or

experience. In “Cultural Differences and Usability Evaluation: Individualistic and Collectivistic Participants Compared,” Hall, Jong, and Steehouder question whether a user’s cultural background affects the usability testing process and results. Hall, Jong, and Steehouder argue that the cultural background of users should be considered a key variable in usability testing conducted with multicultural user groups. Thus, Hall, Jong, and Steehouder conducted a website usability test on users from Western Europe and Asia/Africa and comparing the process as well as the results. Both groups consisted of male PhD students who spoke English as a second language, and all the students tested a Dutch website about science that was written in English. These measures were taken to try to remove gender and education variables from the study while incorporating users of different cultures. Hall, Jong, and Steehouder also used Hofstede’s four item Individualism index, or IDV, to test the participants. Hofstede’s IDV index tests whether an individual or an entire culture is individualistic or collectivistic in nature.

The results of the two user groups showed some differences between how the two groups used the website, most notably in user behavior during the test. However, it seemed that users from both groups found or experienced many of the same problems with the website. When Hofstede’s IDV index was used to interpret the results, Hall, Jong, and Steehouder found that while cultural backgrounds of testers may cause some minor differences in the usability testing process and results, other factors such as gender and education level appear to have a greater affect on usability testing. However Hall, Jong, and Steehouder believe that their study revealed the need for a better understanding of the many factors that may influence the choice of evaluation methods and their results (501). Also, Hall, Jong, and Steehouder question their test results because they feel that other factors such as social desirability (the desire to adjust to another culture, in this case the Dutch culture that created the website) and the selection process

may have obscured the results. Also, they realized that Hofstede's IDV index for evaluating the results may be biased, and a new way to evaluate the test results may be necessary to get more accurate feedback from this study.

Because of the unforeseen variables that may have skewed the results of the study, there is no clear answer for how the website usability test was affected by multicultural users; however it is clear that project teams should keep in mind the cultures of their user groups when planning and conducting usability tests in hopes that their documents and products can be made more user friendly. Also, if a user audience consists of multicultural user groups, the project team may need to select users from each distinct cultural group for the usability test in order to get feedback from each culture and thus make the documentation or product user friendly to a wider audience.

Another factor when choosing users for the testing process is the users' knowledge of the type of product being tested. In "Cross-User Analysis: Benefits of Skill Level Comparison in Usability Testing," Laura Faulkner and David Wick studied the results of usability tests performed by an equal number of novice users and expert users. Their study revealed that the novice users find usability issues concerning what Faulkner and Wick call the "learnability" of the product or document, while the expert group finds issues concerning the efficiency or "optimal use" of the product (Faulkner and Wick). Oftentimes project teams neglect to test a product on expert users, which can cause avoidable problems once the product or documentation is completed and released to the public. One grave example that Faulkner and Wick give concerns a chemotherapy machine that killed or severely injured six patients. Doctors who were experts at entering the client's treatment information into the machine were quicker at entering the information into the machine than the machine was at processing it, thus resulting in the

death and injuries of the six patients. One doctor, after an accident with a patient occurred, actually tested the machine by rapidly entering information and discovered the cause of the usability problem. (Faulkner and Wick) These accidents were avoidable; the machine's reading error would clearly have been discovered through conducting tests that included expert users. Based on this example, and some more theoretical studies that Faulkner and Wick conducted, it is clear that project teams should test not only novice users but also expert users in order to discover more of the usability problems that exist with the product or documentation.

Faulkner and Wick advise a pre-screening process for users during selection for testing so that an equal number of novice users and expert users can be included in the test. Essentially, novices and experts are two distinct user groups, and as Caulton suggested, users from each distinct user group should be tested in order to find a higher percentage of the usability problems.

For technical communicators who have strict budget or time constraints, it would be most efficient to identify distinct user groups based on knowledge and tasks to be performed. For example, using the banking computer program, two distinct groups would be IT personnel who install the program, and bank tellers who use the program. From within those two groups, novice users and expert users could also be identified resulting in two distinct user groups with two distinct user sub-groups. To account for any differences that may result from age or cultural background, the technical communicators could try to recruit users from the distinct user groups who represent a variety of age groups and cultural backgrounds.

Finding Test Subjects that Represent the Users

In Chapter Three I explained that companies with larger budgets would benefit from testing a larger group of users, preferably at least five users from each distinct user group that is

identified. I also explained that if a company has a small budget and cannot afford to test a large number of users then testing a smaller number of users is worthwhile. I suggested following a model such as Virzi's model of testing only five users total, or Geoff Hart's advice to test at least one user because both Virzi's and Hart's options result in identification of usability issues that could affect all users. Therefore, in this section I focus on how to find and recruit users for the test and how to compensate the users for their time.

There are two options for technical communicators to recruit users. The first option is to use a professional recruiter. Professional recruiters have databases filled with a wide variety of people and are likely to have candidates who represent the distinct user groups that technical communicators have identified (Kuniavsky 87). Mike Kuniavsky advises technical communicators to provide the recruiter with a complete user description and screening questions (115). Providing the recruiter with this information will help the recruiter find quality users for the test.

The second option technical communicators have is to build their own database. Mike Kuniavsky advises technical communicators to begin by recruiting friends, family, neighbors, employees of neighboring businesses, and new co-workers who represent the product and documentation user groups (88). As the technical communicators' experience with the recruitment process grows Kuniavsky recommends expanding the user database using the following methods: community emailing lists, online bulletin boards, referrals from past research participants, ads, and community bulletin boards (91-2). He also advises technical communicators to check with the sales and marketing departments within their companies to see if they have a database of users who may be appropriate for the usability test (92). The recruitment process can be very time consuming and difficult for technical communicators to

juggle in between their other work requirements. In most cases one person will need to focus all of his or her efforts on recruitment during the usability testing phase of the documentation project. To reduce costs Kuniavsky recommends training a temp or an intern to recruit full-time (84). At a company with a larger budget or a full-time usability testing team a member of the team can be given a permanent position as recruiter and can maintain the user database.

Before recruiting users as test subjects, technical communicators must know how they will compensate the users for taking the time to perform the test. If using a professional recruiter, the recruiter will charge a price per user that includes compensation for the user and technical communicators do not have to provide the users with any additional compensation (Redish and Dumas 147). Oftentimes users are given money for participating in usability tests but Redish and Dumas also recommend the following incentives, which may be used in lieu of money: gift certificates, travel to the test site, samples of the product, or samples of other products (148). Once a form of compensation is set, technical communicators must remember to include the cost in the budget for the usability test.

Determining the Budget

When determining the budget for usability testing it is important for technical communicators to include “[p]eople’s time (including your time and the research team’s time), recruiting and incentive costs, and equipment costs” (Kuniavsky 75). To estimate people’s time, Kuniavsky recommends using the following rough estimates: ten hours for preparation, two to three hours per user for recruiting and scheduling, three hours per user for usability testing, twelve hours for preparing a report on the results, and six hours for preparing a one-hour presentation based on the report (76). Incentive costs will depend on the amount technical communicators feel is appropriate to compensate users for their time as well as entice users to

participate in the test. Equipment costs depend on the type and amount of equipment technical communicators use for the test, and as previously mentioned in Chapter 3 can be as simple as a single workstation, a stopwatch, pen, and paper or as elaborate as a fully equipped laboratory complete with computer workstations, video cameras, and a viewing room with one-way glass.

For technical communicators who do not have a large budget available to them for usability testing, Hart suggests introducing the concept of usability testing by first conducting many small, low-budget tests, to pave the way for larger scale tests. Starting with smaller tests gives technical communicators experience with testing that can prepare them for more complicated testing, and by starting small technical communicators can prove the value of the tests to supervisors and thus secure larger budgets and resources for testing in future projects. Hart's myth-busting information provides a great starting point for technical communicators who have never conducted usability tests before but would like to incorporate them. It also provides options for technical communicators who have an inadequate budget for a specific project and cannot afford a more extensive usability testing process.

Choosing a Testing Procedure that Fits the Budget

Another consideration when planning a usability test is to figure out which testing method should be used. Monique Jaspers discusses the advantages and disadvantages of three popular usability testing methods: the heuristic evaluation, the cognitive walkthrough, and the think aloud in her article, "A Comparison of Usability Methods for Testing Interactive Health Technologies: Methodological Aspects and Empirical Evidence." She also identifies situations when one method has an advantage over another method and concludes that no one method is consistently best in all situations. Ultimately, Jaspers' advice to usability testers is to use more

than one method and to choose methods that complement each other so that more usability problems can be found and corrected.

Determining the Schedule

To determine the schedule, technical communicators must see whether they can plan the usability testing alongside the product development or whether the technical documents will be written after the product development has been completed. Technical communicators will need to decide how long the usability testing process will take. Many of the scheduling decisions will depend on the budget and the product and documentation development schedules.

There are two main ways to determine the schedule for usability testing. One is to consult with product team manager to schedule testing alongside product development (Kunsiavsky 65). Joann Hackos advises technical communicators to create prototype technical documents during the early stages of product development and to update the prototypes as new information about the product becomes available, which then gives technical communicators the opportunity to schedule usability testing on the prototype document throughout the product development cycle (439). If technical communicators must write the documentation after product is completed or is nearly completed and past the product testing phase, technical communicators will need to base the document usability testing schedule on the documentation project schedule and may be forced to complete testing at the end of the documentation project, which leaves less time to revise the documents based on test results. When at all possible, technical communicators are encouraged to schedule usability testing earlier in the product and document development schedule to allow time for changes in the product and the documents based on usability testing results. Results from any usability testing done at the end of the

development schedule can be recorded and applied to the next release of the product and technical documents (439).

A multitude of factors can influence the amount of time needed for testing processes, many of which are discussed in by Joseph Dumas and Janice Redish in their book *A Practical Guide to Usability Testing*. Some examples of factors that can influence the time usability testing takes are the amount of usability planning that has already been done on the project, the complexity of the project, the extent the product will be tested, how many users are necessary for the test and how much time each user needs to spend on the test, and how many other tasks the project team is doing at the same time (101). These factors vary from company to company and from project to project and must be assessed by each project team in order to estimate the amount of time that will be necessary for a usability test. Dumas and Redish found that for many companies, the entire usability testing process can take eight to twelve weeks, but can be shortened to four to six weeks (101-2). In companies that are testing small sections of the product features or the documentation and that have experienced staff using established usability testing protocols and testing facilities, usability testing can be accomplished in as little as one week (103-4). Dumas and Redish encourage technical communicators with less usability testing experience and fewer resources to aim for anywhere from four weeks to twelve weeks for the testing, keeping in mind that progress on the project and documentation continues during this time period. However, if a company is unwilling to provide adequate time for usability testing, Dumas and Redish encourage technical communicators to perform the testing in whatever time the company does allow. Dumas and Redish say:

You *can* have a valuable test with real users and real tasks in a couple of weeks.

It may mean skimping on how many participants, how many tasks, and how much

analysis you do...[b]ut if it convinces managers that usability testing is worthwhile, it may mean that you will get the resources you need to do more in the future. (103)

Thus, when scheduling for usability testing, technical communicators who are novice usability testers can schedule as few as two weeks but should try to schedule eight to twelve weeks for optimal results. Technical communicators can gradually reduce the amount of time necessary for usability testing as they gain more experience and become more efficient in the entire testing process.

Writing the Test

In order to write the usability test, technical communicators must first set goals for the test, then set measurements for the test, and lastly plan to protect against data loss. Next, technical communicators can write a task-based test for users based on testing goals for the technical documentation.

Setting Goals for the Test

Carol Barnum's chapter on planning for usability testing, from her book *Usability Testing and Research*, explains how to set goals and measurements for the test. The two main questions Barnum poses to readers are what do you want to learn from the test? How do you want to measure the results you receive (153)? The project team must decide which issues are most prevalent and must create a list of well-defined questions that address these issues. Barnum states "Once the team understands its goal for testing, it can turn its attention to the users' goals" (154). The team's questions about the documentation or product become the basis for constructing the usability test. The questions get revised into a list of measurable goals, for

example a goal that the user will perform a task with a maximum of three clicks. According to Barnum, by creating a list of questions and measurable user goals based on the questions, the project team can find many ways to accurately test their document or product and to solve any problems or errors that occur during testing.

Mike Kuniavsky also advises technical communicators to come up with a list of goals for the usability test. He presents three steps to coming up with measurable goals for the test: first, collect issues and present them as goals; second, prioritize the goals; and third, rewrite the goals as questions to be answered (59-62). He further advises that “[t]he mantra for determining which questions to ask is simple. *Test what’s testable*. Don’t ask questions whose answers won’t be actionable or test things that you can’t change” (65).

Don Reinertsen also stresses that usability tests must present testable, measurable questions. In his article, “Don't Define A Requirement Until You Know How To Test For It,” he explains that product specifications are not usually measurable until a test plan is defined (50). As a result, oftentimes products pass usability tests because their specifications were met. However, since the specifications were not measurable, the product’s usability has not actually been ensured and problems often still exist. The easiest way to avoid this problem is to only include requirements in the product specifications that are measurable—Reinertsen suggests asking the question “How will we test this” (50)?

Setting Measurements for the Test

Once goals and measurable questions have been set for the test, technical communicators must set measurements for the test. In “Character-Based, Task-Based, and Results-Based,” Saul Carliner provides detailed information on a task-based approach to assessing usability. He

explains that “For criterion-referenced assessment to be effective, people must begin the design process by defining observable and measurable objectives for the communication product” (Carliner 91). This concept applies directly to usability testing—in order to glean useful information for improving documentation or products, the project team must construct a test that consists of several specific and measurable objectives or tasks for the user to complete that can show the project team where there may be problems in the documentation or product. Two examples are overly complicated instructions for a task or an error resulting from an incomplete line of code.

One simple way to achieve a measurable outcome is to test small sections of the documentation or product in each test rather than testing the entire product in one test. By testing small sections, the project team can assign the user a specific set of tasks to perform that will pinpoint trouble spots in the product. Another way to incorporate Carliner’s task-assessment concept into usability testing is to construct unbiased feedback questions for the user to answer. To make the feedback questions unbiased and thus more measurable, the writer must keep the questions neutral by omitting language that could influence the user’s answer, such as, “What is wrong with this section of the manual?” Asking the user to point out what is wrong with the manual will force the user to come up with problems even if he or she did not actually experience any problems with that particular section.

Overall, Carliner’s task-based approach to usability testing provides a simple template for any type of usability test and coincides with Barnum’s approach to preparing the tests. Without measurable questions and tasks, usability testing will not produce fruitful feedback for the project team.

Brinkman et al.'s findings in their article "Component-Specific Usability Testing" also support smaller scale testing. Brinkman et al. have determined that fewer participants are required in usability tests of different versions of a component due to an increase in test effectiveness when component-specific usability tests are used. The increase in test effectiveness is due to the limited scope of the test. The limited scope of the test means any usability problems that are found can be pinpointed to the specific component that is being tested. This makes it easier for the company fix usability problems because it saves the company time and money that would have been spent searching for the faulty component. This information could be adapted to testing technical documents to help technical communicators save money on usability testing through reducing the number of test participants and pinpointing the cause of usability issues more accurately by testing smaller sections of the usability document at a time, perhaps one task at a time rather than an entire chapter at a time. Again, these findings complement Carliner's by showing that testing small sections of products or of technical documentation can be more effective than testing an entire product or technical document at once because problems can be pinpointed to specific components of a product or specific sections of a document.

Protecting Against Data Loss

Part of the usability test planning process includes planning ways to protect against data loss. While recording observations during usability tests, there are many ways to inadvertently introduce data loss. Spencer's chapter on data and reporting in his book, *Computer Usability Testing and Evaluation*, discusses visual observations and manual recording. Spencer also discusses the limitations of eliminating data loss due to a variety of causes such as interruptions

in the observation or recording process, limitations to the brain's ability to process information, and splitting attention between more than one test subject. Spencer states "Knowing those limitations can help you to plan data gathering and reporting so as to avoid data loss...testers must plan so that they know what they are looking for, and will know what constitutes an aberration from the norm" (127). When planning a test, the project team should limit data loss by controlling the causes of it when possible. For example, the project team could assign one observer per user so that data loss from splitting attention between multiple users could be prevented. Another example is that the project team could have the test in a laboratory or other controlled environment to prevent interruptions during the test. In addition, the project team could use equipment such as video cameras or voice recorders in order to reduce data loss due to human error. Data loss prevention is an important element to consider when planning a usability test and the data loss prevention options that technical communicators will have will depend on budget and time constraints.

Conducting Usability Tests

The two most common usability testing methods for technical communicators are field tests and laboratory tests. Occasionally surveys are also used for usability testing. In this section I will discuss the benefits, limitations, and implementation of field tests and of laboratory tests. I also briefly discuss using surveys for usability testing by sharing two examples of companies that used this approach. I finish by discussing how technical communicators can create a usability team.

Conducting Field Tests

When conducting a field test, the project team sends one or more members out to observe users testing the product in their own workspace, often at an office. In this section I discuss the benefits and limitations of field tests, and how to implement field tests.

Benefits

Field testing can produce more accurate results than laboratory testing because the users are observed in their own work environment, which means the project team can see how the users will interact with the document or product on a day-to-day basis. The project team also gets to see the effect of environmental factors on the usability of the product, such as the size of the work space, the types of sounds that may interfere with day-to-day work, and the quality of the technical resources available to the user. Barker claims that field tests are also less expensive than laboratory tests, mostly because the project team is spared the expenses of building a lab and purchasing equipment (240). Some other benefits of field testing are:

1. You go to the user, rather than asking the user to come to you,
2. You get to observe the actual environment in which the user works or lives, rather than attempting to simulate it artificially,
3. You get to study the artifacts in the user's environment, such as sticky notes on the user's monitor, personal job aids, and other techniques created by the user to remember certain tasks or functions, and
4. You learn whether the user has access to documentation and where it is located (and used). (Barnum 20)

Limitations

There are limitations to field tests such as data loss or skewed results from interruptions, distractions, inadequate materials for the test, and an overall lack of control over certain test variables. Errors can also occur because the users feels as though “someone is looking over their shoulder” and are made uncomfortable by this outside presence in their work area (Barker 241). This discomfort can cause the user to perform tasks differently than he or she usually would or can distract the user and result in more errors or overall difficulty during the test. Some other limitations of field tests are:

1. You typically cannot test as many users, as companies may be unwilling to allow frequent disruptions (not to mention the additional time and expense required for the product developers to be out of the office), and
2. Because field testing is usually a summative evaluation, complex problems uncovered during the testing generally cannot be changed on the product tested.

(Barnum 20)

Implementation

The best way to implement field tests so that the information from the test can be used in a timely manner to improve the product and product documentation is to schedule structured observations. Structured observations consist of a user being asked to perform the kinds of tasks that might also be done in a lab, such as completing specific tasks or using a manual to resolve a problem. The other way field testing is conducted is through longitudinal field studies, which consist of users being observed periodically over a longer time frame to see how they progress as they gain experience with the product. Longitudinal field studies are very costly (they take more

time and money to complete) and the results are often not available in time for the first release of a product or product documentation, thus making it an unsuitable option for usability testing. Structured observations are much better suited to the time and budget constraints that technical communicators face during the documentation production cycle. Structured observations are scheduled for a specific date and time and technical communicators can assign specific tasks for the user to complete, just like during a laboratory test. This allows technical communicators to see how a user would use the product in his or her actual environment. The environment can make a difference in the usability of the product. The structured observation technique can be used earlier in the product cycle or at the end. When this technique is used at the end of the product development cycle, the results are not used to improve the current product but in a later release. (Barnum 21)

To implement field testing, technical communicators must make arrangements with a company that is willing to allow the technical communicators to interrupt their employees' day with testing. The best candidates for field testing are companies that use the product or plan to use the product because they will have more incentive to participate in the field testing—it gives the company a chance to scope out the product in advance and have input on the final design. The marketing department may be the best source for finding a company that will be willing to participate in the field test; some of the techniques for finding users from the section in this chapter titled “Finding Users Who Represent the Audience” may also be used to find companies for field testing. Compensation for the time the employees, or users, are taken away from their job may not be necessary if the company feels that getting a sneak peek at the product is fair compensation for their employees' time.

The final step in implementing field testing is to make travel arrangements for technical communicators to the company where the field testing will take place. This may include plane tickets and hotel rooms or could be limited to reimbursement for gas depending on how far the technical communicators will be traveling to reach the company. Technical communicators must also consider what tools or materials they must bring with them to the field test location while making travel arrangements. The tools or materials may include the product, hard copies of the product documentation or electronic storage devices if the documentation is not a hard copy, tape recorders or video cameras, or pens and notebooks.

Conducting Laboratory Tests

When conducting laboratory tests, the project team creates their own testing space and provides all the necessary equipment, then arranges for the users to come to the lab to take the usability test. In this section I discuss the benefits and limitations of laboratory tests and how to implement laboratory tests.

Benefits

By providing the test space, the project team can exert more control over testing variables, especially equipment issues, interruptions, and other distractions. They can also devise less obtrusive ways to observe the users such as one-way mirrors or hidden video cameras. The increased control over testing variables can make laboratory tests more accurate than field tests; however this is not always the case (Barker 260). With laboratory testing, users are removed from the environment where they would use the product or documentation, which means that laboratory results may not reveal usability problems that are related to the users' environment.

Limitations

The main limitation of laboratory tests is that the users may feel uncomfortable being in the new environment, which can distract them and result in more errors when performing tasks or in an increase in the time it takes to perform tasks. Also, laboratory tests are typically more expensive than field tests because of the cost of creating and maintaining the testing facility and equipment (Barker 263). It is important to remember that the cost may be significantly less at a company that already has an equipped testing facility that is shared with various departments, so for some project managers the cost of a lab test may not be an issue.

Implementation

First, technical communicators must have a laboratory space. This can consist of a full scale lab or a single workstation with chairs for observers. There are many ways to set up a full scale lab but the most common way is to have two rooms divided by a one-way glass wall. One room is full of computer work stations with dividers between each station so the users have their own work space. The observers watch the users from behind the one-way glass in the second room. Sometimes a third room is set up where executives or other company employees may watch the test via a closed-circuit video camera system. (Barnum 12-3). Companies with less space may choose to set up a single room with work stations for users and have the observers sit in the room during the test to observe and record the users' actions. In both set-ups it is very helpful to have cameras and microphones with recorders in the room to record users' actions and comments during the test in case the observers miss crucial information.

For companies who cannot afford to set up a full scale lab, another option is to set up a temporary lab using a small room with one or two workstations and some chairs for observers.

Video cameras and microphones with recorders are optional if the budget is tight. Instead, the observers can simply record all observations on paper or note cards. Carol Barnum suggests observers use index cards and record one observation per index card (18). A simple stop watch can be used for timed tasks (18).

Survey-Based Usability Testing

In their article, “The Value of Usability Testing in Document Design,” Anita Wright and Jone Rymer describe the usability testing process they followed to test the IRS Form 2119 (used to report income on the sale of a house). The form was short and confusing to tax practitioners as well as taxpayers. When conducting the usability test, Wright and Rymer tested taxpayers and tax practitioners in order to find the usability problems that both user groups experienced with the form.

To get feedback from tax practitioners, Wright and Rymer relied on telephone surveys with ten practitioners. Afterwards, they compared their results and found that over seventy percent of the tax practitioners were making mistakes on a line that affects the results of four subsequent questions, which meant that by making one mistake the tax practitioners were giving incorrect value answers for the majority of the form.

To get feedback from the tax payers, Wright and Rymer conducted individual tests with the tax payers. For the test, they requested that tax payers use the think-aloud method so that they could understand the thought process behind each tax payer’s actions while filling out the forms. Each tax payer had a scenario that gave him or her the information he or she needed to fill out the form and that provided Wright and Rymer with a way of evaluating the test results against the correct results for the given scenario. After the tax payers finished filling out their

forms, Wright and Rymer conducted a structured interview with each tax payer, and thus found which lines the tax payers were having trouble understanding.

Wright and Rymer followed testing procedures that enabled them to gather useful information about the usability problems with the Form 2119 and to pinpoint the exact lines and wording that was confusing the user groups. Once they understood which lines and phrases were confusing to the user groups, they were able to understand why it was confusing and how the users interpreted the confusing lines. For other small documentation or product usability tests, Wright and Rymer's testing approach may also be an effective way of finding trouble spots.

Surveys are not always a successful usability testing method, as demonstrated in Sheri Rosen's article "Actions Speak Louder than Words." In the article, Rosen interviews one company about the company's recent successful overhaul of their intranet. The interview took place after the company used a large usability test based on observation to reveal the real problems that users were experiencing. Originally the company tried using surveys to find usability problems, but the survey results showed people's opinion about items rather than revealing usability problems and causes of usability problems. Since the surveys were not helping to communicate the problems that users were experiencing with the intranet, the project team decided to assign one to two observers per user to record the users' interaction with the intranet while completing the test. The resulting notes quickly revealed that the intranet was poorly organized and had an inadequate search engine that most users did not even know existed (page reference).

Based on these usability issues, the project team was able to better organize the intranet through better indexes and a cleaner homepage with links to all the departmental sites and improve the search engine. An additional observational usability test after the intranet overhaul

showed that users were having a much easier time navigating the sites and finding desired information. This company's experience shows that user-observation testing methods produce more measurable test results than user feedback based methods such as surveys.

A key distinction between surveys and field or laboratory testing is that surveys, questionnaires, and interviews are effective ways to learn about users and the tasks they perform, but that usability testing is a more effective method for technical communicators to learn how users interact with products and product documentation (Hackos 440-1). For this reason, most technical communicators should plan to use either field testing or laboratory testing for their documentation usability testing.

Creating a Usability Team

In order to plan and conduct usability testing it is beneficial for technical communicators to create a usability team. The team may consist of technical communicators who are dedicated to usability testing full time or the team may consist of technical communicators with other primary job responsibilities who help with usability testing as mentors to co-workers with less testing experience or who participate in it as part of their documentation product schedule.

Creating a Full Time Usability Team

As mentioned briefly in Chapter Three, Postava-Davignon et al. wrote about the experiences that a growing company had with usability testing in their article "Incorporating Usability Testing into the Documentation Process." With the company's growth came a need for more usability testing. Originally each writing team in the technical communication department was responsible for handling the usability testing of the team's documents; however the increased workload that came with the company's growth made it difficult for the writing teams

to perform the testing. Documentation team members in this company found that they needed a way to better manage the usability testing process; their solution was to create a full-time documentation usability team. The team consisted of a group of technical communicators with various types of usability testing experience and one usability testing expert. The team helped technical communicators in the department with usability testing, refines testing procedures, and provides a database that tracks results and strategies from various tests and is available to the entire department (Postava-Davignon et al.). Also, the team has created a guide to usability testing for the entire writing department.

Another example of a company that created a full time usability testing team is Ajax Industries. The company's usability testing department began when an employee from the marketing department learned about usability testing from a colleague and wanted to establish usability testing at Ajax. She ran a small scale test on her own while still performing her regular job duties. The test was such a success that she was authorized to perform a larger scale test, and a couple of years later she was authorized to build a testing lab and recruit staff to run the tests. Due to this employee's efforts, usability testing became a regular part of the development process at Ajax and a full time usability team was established. (Dumas and Redish 90-1)

A full time usability team would handle the entire testing process for all technical communicators' documentation projects from recruiting users to helping technical communicators develop testing goals and questions to running the usability lab and tests. Having a full time usability team can help technical communicators spend more time on the other phases of documentation production. However, if a full time usability team is not approved or is out of the company budget there are ways technical communicators can create a part time usability team.

Creating a Part Time Usability Team

Since a smaller company may not have the available staff to commit full-time status to usability testing, technical communicators could nominate a group of employees who have more experience with usability testing to act as usability testing mentors to co-workers and project teams. The mentors could help train staff about usability testing through one-on-one work with staff members or by hosting staff meetings to teach usability testing and to discuss recent usability testing experiences and implications. By incorporating a lower budget form of a documentation usability team, any department can slowly work their way towards a more advanced usability testing process, and in the meantime can greatly improve their current usability testing practices.

Another way to set up a part time usability team would be for all technical communicators to participate in the usability testing phase of their documentation project. Usability testing would be included in the documentation production schedule and specific tasks could be assigned to each technical communicator involved in the documentation project. The documentation project manager would lead the team of technical communicators through the usability testing process. With this set up, all technical communicators would gain experience with usability testing and if at a later date the company decided to set up a full time usability testing team many of the technical communicators would be eligible to work on the team.

Managing the Usability Testing Process

One major tool that can help technical communicators to manage the usability testing process is to create a test plan. The test plan can include the goals and objectives for the test, the user profile, the testing method (field test or laboratory test, equipment for users and observers, placement of observers during the test, and other relevant information), the task list (for tasks the

users will perform during the test), the evaluation methods that will be used, and the dates the tests will be held (Barnum 171).

Another way to help manage the testing process is to assign one technical communicator with the role of being the facilitator on the testing day. The facilitator's job is to greet the users, get them set up at their work stations, brief them on the test using a prepared script, and provide help if necessary. The brief should include a description of the product that is being tested, an explanation of how the room and work stations are set up and where the observers will be during the test, explanation of testing protocols (for example, if the user is expected to use the think-aloud protocol during the test), and an explanation of the testing process (such as asking users to follow specific task instructions or read a scenario and make their own decisions about what to do based on the scenario). (Barnum 234)

When managing usability testing, it is important to be aware of the possibility of going overboard or introducing bias into the testing process. Darlene Fichter addresses some common mistakes, or "pitfalls," that beginner usability testing managers make and shares some warning signs so that these behaviors can be identified and avoided. Two of the pitfalls Fichter mentions are ignoring the evidence and over testing. Testing managers often ignore the evidence of a usability issue by explaining it away or explaining away the user's behavior (Fichter). Fichter advises that whenever it is tempting to explain away test results that project team needs to stop and assess what is really going on. *Over testing*—testing without a clear purpose—is a pitfall because tests should be carefully planned out so that extraneous information is filtered out; project teams should always have a purpose for what and why they are testing. Creating a test plan would help technical communicators avoid the problems that Fichter addresses by helping technical communicators stay focused on their goals for the test.

Evaluating Usability Testing Results

Mike Kuniavsky identifies a three-stage process for evaluating usability test results: collecting observations, organizing observations, and extracting trends from the observations (293). Technical communicators collect the observations during the test using various forms of media such as video cameras, voice recorders, and notes written on index cards or paper. Once the test is complete, technical communicators will review all the media and compile a list of all observations. Then technical communicators must organize the observations by reading through the list to find repetition and issues that are caused by the same main problem (296).

Hackos's chapter on usability testing in her book, *Managing Your Documentation Projects*, gives helpful advice on interpreting the test results to identify usability problems. She points out the fact that oftentimes when a user finds a small problem with the documentation it often is the result of a larger, global problem that the user does not verbalize or often may not recognize. Hackos states that "By correcting the immediate problem, you may neglect the root cause of the problem, which means, in all probability, that the same type of problem will occur again" (542). Thus, it is up to the project team to go beyond the small, quick fix to see if any larger problems exist that would need to be addressed.

The example that Hackos provides is of a user's complaint that she could not find something in the index. The user included the list of terms she had searched for while trying to find this information. The quick fix would have been to include one or all of these terms into the index. However the project manager found that the larger, global problem was that the writing team had only indexed the headings, not the terms within the text. (543) Therefore, for every problem that comes up in the usability test it is imperative that the project team carefully

evaluate all probable causes to come up with the best solution rather than settling for the easiest solution.

Dumas and Redish agree with Hackos that the scope and severity of the problem need to be assessed, which means finding out whether the problem is global or local. Dumas and Redish define scope as meaning “how widespread is the problem?” and severity as meaning “how critical is the problem?” (322). While finding out whether a problem is global (wide scope, may effect large portions of the document and cause many instances of a usability issue) or local (narrow scope, only effects specific sections of the manual and may only cause one instance of a usability issue) is important, Dumas and Redish also stress the need to organize usability issues by level of severity because a local issue may have more severe consequences than a global one (324-5). Dumas and Redish describe four levels of severity for usability issues. Level 1 problems prevent completion of a task, Level 2 problems create significant delay and frustration, Level 3 problems have a minor effect on usability, Level 4 problems are subtle and tend to be enhancements that can be added in the future to increase usability (324-5).

Using the scale Dumas and Redish present would help with problems that professionals in the technical communication field face when judging the severity of usability problems with documentation that Lentz and De Jong discuss in “How Do Experts Assess Usability Problems? An Empirical Analysis of Cognitive Shortcuts.” There are several shortcuts that experts inadvertently take when guessing how novices will react to a document or a task that are based on the experts’ own experience or knowledge. The result is that experts do not always agree on the severity of a usability problem and can overlook or explain away valid problems that users have with a document. Lentz and De Jong explain that when experts are faced with judging the severity of a usability problem, they need to determine whether the problem impairs the readers’

understanding of the document (or section if testing smaller portions at a time), and whether this misunderstanding prevents the reader from getting the primary information they need from the document or if it prevents the authors of the document from achieving their primary objective. If so, then the problem is serious and must be addressed in the revisions.

When technical communicators are compiling test results and identifying usability problems, it is useful to keep in mind Hornbaek and Frokjaer's studies on the "evaluator effect," which is the observation that usability observers in similar situations often identify very different usability issues, reveal that using a group of evaluators to interpret usability test results can increase the effectiveness of the test in some ways. They found that matching answers between the different evaluators and grouping together similar findings results in a shorter list of unique problems. The more usability problems that evaluators can agree on, the more likely the problems will be resolved before the final product is shipped. Hornbaek and Frokjaer advise practitioners to use teams of evaluators to analyze usability problems and interpret usability testing results and to group similar answers together to get a more comprehensive view of the usability of the product or document.

Once technical communicators have completed the three-stage process for evaluating test results, it is important to write a report of the findings. The report will help technical communicators work with management to determine which issues need to be resolved based on the severity (such as using Dumas and Redish's levels of severity) and how much room there is in the budget and schedule to address the issues. In "Usability Testing in Technology Education," Jim Flowers provides advice that would be helpful to project teams about how to format the report of the test results. Flowers suggests that the report include all of the test results that revealed a problem with the product or documentation, including interview and survey

responses along with observations of users during the test. Next, Flowers says that every problem that is brought up in the testing needs a list of probable causes and a solution, thus the report would include the solutions that the project team created for each usability problem. Flowers also says that the report should contain information about “observed usability problems and the design features of the interface, noting what was found to be unintuitive and why” (18).

To further organize the report, Dumas and Redish suggest formatting so that each problem or issue is clearly identified with its own section and descriptive heading along with a paragraph explaining the problem or issue, with sub-sections that include the scope, the severity level, the frequency, explanation that can include feedback from users or user observations that alerted technical communicators to the problem/issue or some examples of the problem/issue, and recommendations to resolve the problem/issue. Making each problem or issue its own section with descriptive headings makes it easy for others in the company to use the report itself as a checklist while resolving each issue. Also, managers or others reading the report may prefer to skim the report or skip around to issues that concern them the most rather than reading the report straight through. Formatting the report so that others can easily see each problem/issue and the recommendations increases the chances that the usability issues will be resolved. (350-1)

Evaluating Usability Testing Methods

Some project teams may try to find ways to simplify the usability testing process by relying on past test results to apply to current projects. In “What Do We Manage? A Survey of the Management Portfolios of Large Technical Communication Groups,” Saul Carliner briefly discusses how some technical communicators have attempted to apply the results from one usability test of a document to all the documentation they write rather than conducting a usability test for each separate documentation project. Some authors think that the insights gleaned from

testing one documentation project can apply to all documentation projects. These authors create a list of heuristics and principles of user-centered design based on those limited testing results then apply these heuristics and principles to all future projects without even assessing whether these heuristics and principles really work (Carliner).

In his article, Carliner explains that the attempt to create a usability template that works for every project based off of the insights gained from one particular project is faulty because the project team will never find out if their template actually works for the projects they apply it to. Audiences and their needs can change, and new projects can have different errors or usability issues than previous ones did. These differences in audience, audience needs, and projects can create new usability challenges that would go unnoticed if the usability testing process were skipped. Thus, it is important to avoid trying to streamline the process by relying on old test results for different projects; instead, a usability test should be administered for each project in order to account for the inherent differences between different projects and their various user audiences. However, technical communicators should evaluate each usability test from beginning to end and determine how the process can be improved for future tests.

Conclusion

In this chapter I discussed ways for large companies or companies with large budgets, and small companies or companies with small budgets to plan, conduct, and manage usability tests and evaluate usability test results and methods. In each stage of the testing process I presented ways that technical communicators can adapt the usability testing process to fit their companies' budget and resources. I also shared ways that technical communicators can gradually develop usability testing at their company from testing one user while recording observations on note cards to running a full time usability lab with the latest equipment. There is

ample information in the chapter for technical communicators to plan and implement usability testing successfully as well as evaluate and use the results from the test.

CHAPTER FIVE: CONCLUSION

Any size company on any budget can afford usability testing. In fact, technical communicators cannot afford not to perform usability testing. Users are demanding that products and their technical documentation be usable. Companies such as Boeing Co. have decided that new software purchased by the company must have gone through usability testing because products that have not gone through usability testing have wasted company time and money (Thibodeau 1). Experts in usability testing and technical communication have identified a need for usability testing of technical documentation. Barnum explains that “[o]ut of market necessity, resulting from changes in the user population and the increased competitiveness of the marketplace, has come a need for usable products with the resulting need for usability testing as a key element” (3). Users are no longer content to fumble around with difficult manuals and confusing help menus. More technical communicators are realizing the importance of usability testing, as evidenced by the existence of a usability and user experience group within the Society of Technical Communicators (<http://www.stcsig.org/usability/>), and the increase in articles and other research in the usability testing field.

Despite these trends showing that usability testing is on the rise, there are still a large number of practitioners in the technical communication field who do not implement usability testing. I think part of the problem is that, though much of the research I found discusses the cost effectiveness of usability testing, none of the research puts cost effectiveness in the forefront, and I think that the main concern of companies is cost effectiveness. While everyone wants to provide quality service and products to their customers they also want to make a profit, and both time and money are limited resources and must be spent wisely. In addition, many technical communicators think that usability testing is too complicated. In “Ten Technical

Communication Myths,” Hart acknowledges these concerns about usability testing and says “usability testing need not be prohibitively expensive and logistically difficult...[t]he fact that so many of us have been scared away from even considering usability testing is unfortunate” (297). He addresses the concern about making usability testing cost effective and easier to implement by showing that usability testing is cost effective and including examples of ways that companies with small budgets can incorporate usability testing. Hart also provides information on each stage of the usability testing process to help technical communicators immediately begin implementing usability testing at their companies.

Contribution to the Technical Communication Field

This thesis investigates the reasons why there are technical communicators who do not implement usability testing of their documents and determines that the main reasons are because technical communicators and their companies feel that usability testing is too expensive and complicated. This thesis refutes them by discussing the cost effectiveness of usability testing and sharing ways that companies with large budgets and companies with small budgets can begin incorporating usability testing. Next, information is provided on all the steps that are necessary for technical communicators to implement usability testing of technical documentation at their company. Options are presented for everything from bare minimum usability testing with a shoe-string budget with pencils, note pads, and only a handful of users to full scale usability testing in large laboratories with the latest equipment and a wide variety of users. The research provides examples from real companies, advice from experienced technical communicators and usability experts, and research demonstrating how many resources are truly required to benefit from usability testing. The goal of this thesis is to show technical communicators that usability testing is cost effective and that there are many options for implementing usability testing no

matter how large or small their budget is, and that these options can be implemented immediately, thereby empowering technical communicators to start including usability testing as part of the documentation production cycle at their companies.

Once technical communicators are able to include usability testing of technical documents at their companies, they can become user advocates for groups of users who are often misrepresented or underrepresented. First, technical communicators must identify user groups that are misrepresented or underrepresented. Caulton's theory, which states that usability testing results are more accurate if distinct user groups are identified, can also be used to identify groups of users that are typically underrepresented or misrepresented. However, Caulton himself only mentions level of expertise among users as a way of finding distinct user groups; for example he mentions programmers and secretaries as two distinct user groups. Each group would perform different tasks and have different levels of expertise, thus when finding users Caulton suggests finding users from each group, programmers and secretaries, who are beginners, intermediates, and experts rather than finding a random mix. Essentially, he suggests tracking experience as a variable during the usability test (7). Thus, identifying the variables that each distinct user group presents would be important in order to best represent the largest percentage of users. Some common variables that can reveal misrepresented or underrepresented user groups are expertise, disability or physiological differences, and age.

Expertise is a common variable between different user groups and should always be considered when finding users for usability research. In some cases, a difference in expertise level can reveal a misrepresented or underrepresented user group. An example of when expertise reveals a misrepresented user group is in "Keeping Users at the Center: Developing a Multimedia Interface for Informed Consent" by Kim et al. In the article, Kim et al. identify four

distinct user groups of the informed consent instructions they developed for clinical research trials: teenage patients, parents, non-physician caregivers, and physicians (336). Their opinion is that “most systems [for informed consent] tend to privilege the administrative needs of the researchers and the hospital” (335). Furthermore, the premise of Kim et al.’s paper is “that to place the pediatric patient or parent needs anywhere but at the forefront of the design effort is to dilute the intent of informed consent—to protect patients” (336). Thus, while Kim et al. kept the needs of caregivers and physicians in mind during the project, they became user advocates for two distinct user groups they felt were underrepresented, pediatric patients and their parents. For both of these user groups, expertise was a variable that was tracked during the development of the informed consent instructions to ensure that they would understand the informed consent instructions as well as the groups with higher expertise in medical experiments and informed consent. When technical communicators identify a gap in expertise levels between users it is important to perform a similar analysis to see whether there is a user group at a different expertise level that may be underrepresented due to the needs of other user groups so that the inequality can be resolved.

Another variable that can identify a misrepresented or underrepresented user group is disability or physiological difference between user groups. Physiological differences can include differences in size, gender, vision, coordination, or any other physiological variables that would have a large affect on how the user interacts with the product or documentation. For example, in “Helping Low-vision and Other Users with Web Sites That Meet Their Needs: Is One Site for All Feasible?” Mary Frances Theofanos and Janice Redish become user advocates for low-vision users of websites by performing a study to see how low-vision users’ needs differ from those of other website users groups. During their preliminary research, Theofanos and Redish found that

low-vision users of websites are more prevalent than most people think and are expected to double in the next three decades, especially since many common diseases that people develop with age mean that low-vision is a threat to many website users who currently do not have low-vision (10). During their work with the low-vision users, Theofanos and Redish found that low-vision users have web browsing habits that are nearly identical to users without visual impairments (11). During the test, the low-vision users used special software to help make websites more visible to them. Theofanos and Redish observed how the low-vision users adjusted the special software settings to see if there were similarities between how different low-vision users customized the software settings, however they found that “[m]any did extensive customizing...[t]hey changed the level of magnification, the background and/or text colors, the pointer color, the pointer size, the voice on/off setting, the way the screen was magnified, and so on (11). During the study the customizations that the low-vision users made to the special software cause a variety of usability issues for them while viewing various government websites and attempting to perform a list of tasks. In order to be user advocates for underrepresented user groups, technical communicators must identify user groups with physiological differences, such as low-vision users, and determine whether the physiological differences will present significant usability issues for these user groups that need to be accommodated by finding representatives of those user groups to participate in the usability testing.

Age is another variable that can be overlooked when identifying distinct user groups but that could have a significant impact on usability of technical documents for users. Older adults comprise a user group that is often overlooked in technical communication, especially technical communication that is electronic or on the web. The affects that age can have on users’ abilities to perform tasks with products or documentation often overlap with variables of other

underrepresented user groups such as user groups with disabilities/physiological differences, user groups from different cultures, and user groups with a different expertise level (O’Hara 425-432). In addition to those variables, older adults also experience the effects of a prevailing stereotype that older adults are weak, confused technophobes—age-related stereotypes such as these are sometimes referred to as “ageism.” O’Hara states “the American media often promote and reinforce the common stereotypes of old age” (429). As examples, O’Hara cites the “exaggerated hapless Grandpa on *The Simpsons*” and “the woman depicted in the famous ‘I’ve fallen and I can’t get up’ commercial,” which all depict older adults as “weak and confused” (429). Furthermore, O’Hara states that depictions such as these in the American media can have an affect on older adults themselves, thus making it even more difficult for older adults to use newer products and documentation such as websites (429). Technical communicators are in a prime position to be user advocates for older adults via usability testing because they can ensure that older adults are represented in the usability testing so that technical documents can be designed to accommodate older adults from the beginning instead of attempting to make changes later when it becomes more burdensome and expensive. While accommodations for other user groups, such as low-vision or blind users, can also increase usability for segments of the older adult user group, technical communicators will need to consider the affect that ageism has on older users. O’Hara advises technical communicators to ask whether the document provides an ethical and honest attitude towards older adults, whether tutorials are offered to assist older adults, whether the document avoids ageist or stereotypes, and whether the document strives to meet accessibility standards—particularly one that relate to vision, hearing, and mobility (443).

The research explored here about becoming a user advocate indicates that expertise, physiological differences, and age are major variables technical communicators can use to help

identify distinct user groups that are misrepresented or underrepresented and in need of user advocates. Further research on variables that identify a distinct user group and necessitate representation of the user group in usability testing might reveal other variables for technical communicators to consider when striving to be user advocates.

Future Research on the Topic

There are two types of research on usability testing that would be greatly beneficial to technical communicators. The first type is research about how specific companies have integrated usability testing of technical documentation. It would be very helpful to everyone in the field if more practitioners shared the nuts and bolts of how they have established usability testing and what testing methods and equipment they use. In fact, Carliner has identified a scarcity of systematic studies of technical communication management practices (45). He explains that he often sees members of the Society of Technical Communicators asking questions about various aspects of practice management and the answers from other members are usually based on their experience rather than on research or study results. After a review of the available research in the field, including scholarly journal articles, books, and literature reviews, Carliner attributes this trend to the fact that “systematic studies of management practice in technical communication are scarce” (45). Future research gathered from practitioners about how their usability testing process is managed would help others in the field to begin implementing usability testing at their companies. It would also help practitioners with existing usability testing processes to refine their methods or experiment with different approaches, which could help them increase the cost-effectiveness of their usability testing program.

The second type of research in the field that would benefit technical communicators would explore the ways that the addition of usability testing of technical documentation has

affected various companies. The research would consist of practitioners tracking changes in calls to customer service, changes in sales figures, changes in product reviews, and other data that would show whether usability testing was cost effective at their company (Dumas and Redish 14-17, Barnum 22). If a large number of companies participated in the study then technical communicators would have real-life data about the cost effectiveness of usability testing instead of relying on hypothetical examples and testimonials. Currently, the research on usability testing includes some case studies but the studies do not provide actual data and as a result sound more like testimonials instead of data-driven research. For example, Dumas and Redish present brief case studies of Ajax Industries and Omega Industries in which they describe how each company began implementing usability testing. No hard data is provided to show that the usability testing was cost effective; instead the authors mention that each company was encouraged by the results of the testing to expand their usability testing methods (90-1). Kunviasky discusses how a computer product that he worked on was abandoned. He explains that the product was of high quality and very creative, but had “no understanding of its audience” (8). After that experience, his company began performing usability tests but no data is provided to show the losses the company sustained from the abandoned product or to show how incorporating usability testing into the product development cycle for other products improved the company’s profits. Barnum points out that the software development industry has learned it costs 100 times more to make a change in the next release of a product than to perform usability testing and make changes at the beginning of the project’s development, and a scenario of how much time and money one usability issue might cause a user to lose, but again there is no hard data from real life companies or users to show exactly how usability testing saves money (22-3). Providing hard data would make it easier for technical communicators to calculate the true

savings realized by implementing usability testing. Details on how the usability testing is planned and implemented as well as the context of the testing could be included so that various usability testing methods and their results could be compared to each other, which could result in certain methods proving to be better performers in specific testing situations.

One of the barriers to this type of research is that “many organizations consider productivity information to be proprietary and do not let their workers share the information, so comparisons cannot easily be made across organizations and industries” (Carliner 48). Another barrier is that practitioners would have to be willing to track all the variables that show whether usability testing is cost effective at their company. Experts in the field encourage technical communicators to track productivity data to provide proof to management of the cost effectiveness of the department and secure future funding. For example, Dumas and Redish provide one way of calculating the savings resulting from usability testing. Their approach is to “conduct iterative usability tests with the same tasks” (19). Technical communicators are to track the amount of time that users take to perform the tasks in the first test, then improvements are made to the product or documentation, then users perform the same tasks again and technical communicators calculate whether the users took less time to perform the tasks the second time. After that, technical communicators multiply the time by the number of people who will use the product or document, then convert the time saved to dollars saved in terms of the users’ average salaries (19). Barnum encourages technical communicators to compare the number of calls to the help desk or technical support department after the release of a new product that was tested for usability to the number of calls to the help desk after the release of a previous product that was not tested for usability (50). A drop in the number of calls can provide proof to companies of the benefit of usability testing and help technical communicators secure budget and time for

usability testing of technical documents. Tracking productivity data and savings from usability testing would make it easier for companies to see whether they are benefiting from usability testing. If companies became open-minded to the idea of sharing productivity information about usability testing, and technical communicators took the time to track the ways that usability testing has affected their companies through tracking the productivity of the testing, then the entire field could benefit from the valuable information this research would provide.

Future research in both above mentioned areas would also help professors who teach usability testing of technical documents to their students by providing them with more information about how usability testing benefits practitioners. Sometimes there can be a disconnect between the theories that are taught in the classroom where students are given ample time to include usability testing of their technical documents and the harsh realities of the working world where practitioners have to deal with budget and time constraints and perhaps even unsupportive management. If more research based on the experiences of practitioners were included in the field then it would make it easier to show students how to translate classroom theory into real world results. Teaching students how to implement usability testing in the real world can result means the technical communicators of tomorrow will have the tools to incorporate usability testing of technical documentation, thus improving the usability and quality of technical documentation and showing companies the importance of the technical communication field and of usability testing.

Conclusion

Usability testing of technical documentation is cost effective and can be implemented at any size company with any size budget. Even small, simple usability tests performed with one-five sample users and some index cards, pens, and stop watches can help technical

communicators identify usability issues in their documentation that would have a negative affect on the accompanying products' success and the profits of their companies. If further research were conducted that involved gathering hard data from technical communicators' companies to show the cost effectiveness of usability testing as a whole and of specific usability testing methods, then the entire technical communication field would benefit by providing proof of the cost effectiveness of usability testing and improving usability testing methods, thereby convincing more and more companies to incorporate usability testing of their products and documentation.

REFERENCES

- Andre, Terence S., Rex H. Hartson, Robert C. Williges. “Determining the Effectiveness of the Usability Problem Inspector: A Theory-Based Model and Tool for Finding Usability Problems.” *Human Factors* 45.3 (Fall 2003): 455-82. Web. 27 Nov. 2005.
- Barker, Thomas T. “Chapter 8: Conducting Usability Tests.” *Writing Software Documentation: A Task Oriented Approach*. Pearson Education, Inc., 2003. 240-265. Print.
- Barnum, Carol. *Usability Testing and Research*. Pearson Education, Inc., 2002. Print.
- Brinkman, Willem-Paul, Reinder Haakma, and Don G. Bouwhuis. “Component-Specific Usability Testing.” *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans* 38.5 (September 2008): 1143-55. Web. 14 June 2009.
- Carliner, Saul. “Characteristic-Based, Task-Based, and Results-Based: Three Value Systems for Assessing Professionally Produced Technical Communication Products.” *Technical Communication Quarterly* 12.1. (Winter 2003): 83, 18p. Web. 3 Oct. 2005.
- Carliner, Saul. “What Do We Manage? A Survey of the Management Portfolios of Large Technical Communication Groups.” *Technical Communication* Feb 2004: 51.1. 45-67. Web. 3 Oct. 2005.
- Caulton, David A. “Relaxing the Homogeneity Assumption in Usability Testing.” *Behaviour & Information Technology* 20.1 (Jan. 2001): 1-7. Web. 27 Nov. 2005.
- Dumas, Joseph S., Janice C. Redish. *A Practical Guide to Usability Testing*. Albex Publishing Corp., 1993. Print.
- Faulkner, Laura, David Wick. “Cross-User Analysis: Benefits of Skill Level Comparison in Usability Testing.” *Interacting with Computers* 17.6 (Dec 2005): 773-86. Web. 20 Dec. 2005.

- Fichter, Darlene. "Intranet Librarian." *Online 25.2* (Mar/Apr 2001): 78-80. Web. 27 Nov. 2005.
- Flowers, Jim. "Usability Testing in Technology Education." *Technology Teacher* 64.8 (May/Jun. 2005): 17-9. Web. 20 Nov. 2005.
- Hackos, JoAnn. "Evaluating the Publications Product." *Managing Your Documentation Projects*. John Wiley & Sons, Inc., 1994. 528-544. Print.
- Hall, Marinda, Menno De Jong, Michaël Steehouder. "Cultural Differences and Usability Evaluation: Individualistic and Collectivistic Participants Compared." *Technical Communication* 51.4 (Nov. 2004): 489-503. Web. 27 Nov. 2005.
- Hart, Geoff. "Ten Technical Communication Myths." *Technical Communication* 47.3 (Aug 2000): 291, 8p. Web. 3 Oct. 2005.
- Hornbaek, Kasper and Erik Frokjaer. "A Study of the Evaluator Effect in Usability Testing." *Human-Computer Interaction* 23.3 (2008): 251-77. Web. 14 June 2009.
- Jaspers, Monique W.M. "A Comparison of Usability Methods for Testing Interactive Health Technologies: Methodological Aspects and Empirical Evidence." *International Journal of Medical Informatics* 78 (2009): 340-53. Web. 14 June 2009.
- Kim, Loel, Amanda J. Young, and Robert A. Neimeyer. "Keeping Users at the Center: Developing a Multimedia Interface for Informed Consent." *Technical Communication Quarterly* 17.3 (2008): 335-357. Web. 12 Feb. 2010.
- Kuniavsky, Mike. *Observing the User Experience: A Practitioner's Guide to User Research*. Elsevier Science, 2003. Print.

- Lentz, Leo and Menno De Jong. "How Do Experts Assess Usability Problems? An Empirical Analysis of Cognitive Shortcuts." *Technical Communication* 56.2 (May 2009): 111-21. Web. 14 June 2009.
- Maddox, Michael. "The Simulated User." *Industrial Engineer: IE* 37.6 (Jun. 2005): 22. Web. 27 Nov. 2005.
- Nielson, Jakob. "Guerilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier." *Cost-Justifying Usability*. Eds. Randolph G. Bias and Deborah J. Mayhew. Boston: Academic Press, 1994. 242-72. Print.
- O'Hara, Karen. "'Curb Cuts' on the Information Highway: Older Adults and the Internet." *Technical Communication Quarterly* 13.4 (2004): 423-445. Web. 16 Feb. 2010.
- Percival, L.C., S.K. Johnson. "Network Management Software Usability Test Design and Implementation." *IBM Systems Journal* 25.1 (1986): 92-104. Web. 29 Nov. 2005.
- Postava-Davignon, Christi-Anne, Candice Kamachi, Cory Clarke, Gregory Kushmerek, Mary Beth Rettger, Pete Monchamp, Rich Ellis. "Incorporating Usability Testing into the Documentation Process." *Technical Communication* 51.1 (Feb 2004): 36-44. Web. 3 Oct. 2005.
- Reinertsen, Don. "Don't Define A Requirement Until You Know How To Test For It." *Electronic Design* 48.2 (24 Jan 2000): 50. Print.
- Rosen, Sheri. "Actions Speak Louder than Words." *Communication World* 16.6 (Jun/Jul 1999): 60. Web. 27 Nov. 2005.
- Spencer, Richard H. "Chapter 8: Data and Reporting." *Computer Usability Testing and Evaluation*. Prentice-Hall, Inc., 1985. 126-148. Print.

- Theofanos, Mary Frances and Janice Redish. "Helping Low-vision and Other Users with Web Sites That Meet Their Needs: Is One Site for All Feasible?" *Technical Communication* 52.1 (February 2005): 9-20. Web. 16 Feb. 2010.
- Thibodeau, Patrick. "Users Begin to Demand Software Usability Tests." *Computerworld* 36.48 (25 Nov. 2002): 1-2. Web. 26 Nov. 2005.
- Trager, Louis. "Web Site Usability Tests Arrive." *Inter@ctive* 7.5. (7 Feb. 2000): 22. Web. 27 Nov. 2005.
- Van Der Meij, Hans. "Examining the Relationship Between Quality Writing and Quality Reading." *Technical Communication* 47.2 (May 2000): 195-210. Web. 3 Oct. 2005.
- Virzi, Robert A. "Streamlining the Design Process: Running Fewer Subjects." *Proceedings of the Human Factors Society 34th Annual Meeting, Orlando, FL.* (8-12 Oct. 1990): 291-94. Print.
- White, Martin. "Finding Intranet Solutions: Do You Know The Way To San Jose?" *Econtent* 27.½. (Jan/Feb 2004): 19. Web. 27 Nov. 2005.
- Wright, Anita D., Jone Rymer. "The Value of Usability Testing in Document Design." *Bulletin of the Association for Business Communication* 57.4. (Dec. 1994): 48-51. Web. 26 Nov. 2005.