Moving Forward

11

Some of the concepts and approaches we've described may be new to some readers, and maybe even a bit overwhelming at first, so we wanted to distill the key elements of this book into ten steps to help you move forward. These steps will help you establish a culture of usability metrics within your organization. These are lessons we've learned—sometimes the hard way—over the years.

11.1 SELL USABILITY AND THE POWER OF METRICS

One of the most important factors contributing to the success of a usability program is the degree to which management buys into the whole concept of usability. Several techniques can be very helpful in making this happen. First, we recommend that you bring key decision makers into the lab to observe as many sessions as possible. If you don't have a lab, arrange to use a conference room for the day. A screen-sharing application and a conference call can make for a very effective makeshift observation gallery. Nothing speaks louder than observing the user experience firsthand.

Once key decision makers start to see a consistent pattern of results, you won't need to spend much effort convincing them of the need for a design change. But be careful when someone only observes a single usability session. Watching one participant struggle can easily be dismissed as an edge case (e.g., "Our participants will be much smarter than that person!"). Conversely, seeing someone easily fly through the tasks can lead to a false sense of security that there are no usability issues with the design. The power of observation is in consistent patterns of results.

Another excellent way to sell usability is with short video clips. Embedding short video clips into a presentation can make all the difference in selling usability. The most effective way to illustrate a usability issue is by showing short clips of two or three different participants encountering the same problem. Showing reliable patterns is essential. In our experience, participants who are more animated usually make for better clips. Make sure each clip is short—ideally less than 30 seconds. The last thing you want is to lose the power of a clip by dragging it **289**

TIPS FOR GETTING PEOPLE TO OBSERVE USER SESSIONS

- *Provide a place for observing.* Even if it's a remote session, provide a room with projection or a large screen for observers to watch the session as a group. An important part of observing a usability session is the interaction among the observers.
- *Provide food.* For some odd reason, more observers show up when test sessions are scheduled during the lunch hour and food is provided for everyone!
- *Get the sessions on their calendars.* Many people live by their online calendars. If it's not on the calendar, it doesn't happen (for them). Send meeting invitations using your company's scheduling system.
- *Provide information.* Observers need to understand what's going on. Make sure that a session schedule, moderator's guide, and any other relevant information are readily available to the observers, both before and during the sessions.
- *Engage the observers.* Give the observers something to do besides just watching. Provide whiteboards or sticky notes for them to record issues. If there are breaks between sessions, have them do a quick review of the key takeaways from the last session.

---

out too long. Before showing a clip, provide appropriate context about the participant (without revealing any private information) and what he or she is trying to do.

If bringing observers into the lab or putting video clips in front of them doesn't work, try presenting a few key usability metrics. Basic metrics around task success, efficiency, and satisfaction generally work well. Ideally, you'll be able to tie these metrics to return on investment (ROI). For example, if you can show how a redesign will increase ROI or how abandonment rates are higher on your product compared to your competition, you'll get the attention of senior management.

Finding a champion for user experience in upper management is also very important. This person should help you fight for the user experience, get the budget you need for staff and other resources, and convince other senior managers about the importance of the user experience. The more champions you have, the more impact you will have, and ultimately the better the user experience for your products.

11.2 START SMALL AND WORK YOUR WAY UP

Don't try to boil the entire ocean. When deciding which metrics to use for a usability study, start off with something small and manageable. It's critical that you be successful in your first uses of metrics. If you're trying to incorporate metrics in

routine formative testing, start with categorizing types of issues and issue severity. By logging all the issues, you'll have plenty of data to work with. Also, it's easy to collect System Usability Scale (SUS) data at the conclusion of each usability session. It only takes a few minutes to administer the survey, and it can provide valuable data in the long run. As you get comfortable with some of the more basic metrics, you can work your way up the metrics ladder.

A second phase might include some efficiency metrics such as completion times and lostness. Consider some other types of self-reported metrics such as usefulness–awareness gaps or expectations. Also, explore different ways to represent task success, such as through levels of completion. Finally, start to combine multiple metrics into an overall usability metric, or even build your own usability scorecard.

Over time you'll build up a repertoire of different metrics. By starting off small, you'll learn which metrics work for your situation and which don't. You'll learn the advantages and disadvantages of each metric and start to reduce the noise in the data collection process. In our work, it has taken us many years to expand our metrics toolkit to where it is today. So don't worry if you're not collecting all the metrics you want at first; you'll get there eventually. Also, be aware that your audience will have an adjustment period. If your audience is only used to seeing qualitative findings, it may take them a while to get adjusted to seeing metrics. If you throw too much at them too quickly, they may become resistant or think you just got back from math camp.

11.3 MAKE SURE YOU HAVE THE TIME AND MONEY

One of the most frustrating experiences is having the desire to do the job right but no money or time to do it. It's critical to make sure you have the right amount of time and budget before starting any usability study. Be very careful in planning out how much time and money every usability study is going to take. Ultimately, it comes down to a balancing act. On one hand, you don't want to underestimate and have to go back to your business sponsors for more money and be perceived as inefficient (not keeping to the budget) or as poor planners. On the other hand, if you overestimate, you may be perceived as being too expensive. A good way to accurately estimate time and costs is by first considering two general types of usability studies that use metrics.

1. Some usability studies involve adding a few simple metrics to a more qualitative study. The time and budget for these studies should not excessively surpass those if you didn't collect any metrics at all—perhaps up to 10 percent more. For these studies, it helps to build a track record of how much time and money was actually spent on similar types of studies. Then you can adjust up or down,

depending on how easy or difficult the recruiting is going to be, how many days you're testing, any demands to attend meetings as part of the project, and any special requirements in the analysis or presentation. For the most part, such studies are within 25 percent above or below the average estimate.

2. Other studies are more focused on collecting a wide range of metrics. It's harder to estimate time and budget for these kinds of studies. This might be a site-wide benchmark study, competitive benchmark, or online study. These kinds of studies are done less frequently, so it's more difficult to develop historical data to rely on. In addition, the variance is much greater. For these studies, you can break down your time and cost estimates into three distinct phases of any project: test preparation, data collection, and data analysis/presentation. Additional costs might include participant recruiting/compensation and licensing the technology to run online surveys. In our experience, these studies can run anywhere from half the cost of a traditional qualitative lab test up to twice the cost.

One problem you might encounter is running out of time during the analysis stage. Often you're on a very tight timeline and simply don't have the luxury of exploring the data as much as you'd like. In such cases, you can release the results of the study in two phases. In the first phase, give the top-line actionable results. In the second phase, provide a more detailed analysis of the data with follow-up from the first phase.

11.4 PLAN EARLY AND OFTEN

One of the key messages of this book has been the importance of planning ahead when collecting any metrics. The reason we stress this is because it is so tempting to skip, and skipping it usually has a negative outcome. If you go into a usability study not sure which metrics you want to collect and why, you're almost certainly going to be less effective.

Try to think through as many details as you can before the study. The more specific you can be, the better the outcome. For example, if you're collecting task success metrics and completion times, make sure that you define your success criteria and when exactly you'll turn off the clock. Also, think about how you're going to record and analyze the data. Unfortunately, we can't provide a single, comprehensive checklist to plan out every detail well in advance. Every metric and evaluation method requires its own unique set of plans. The best way to build your checklist is through experience.

One technique that has worked well for us has been "reverse engineering" the data. This means sketching out what the data will look like before conducting the study. We usually think of it as the key slides in a presentation. Then we work back

from there to figure out what format the data must be in to create the charts. Next, we start designing the study to yield data in the desired format. This isn't faking the results but rather visualizing what the data might look like. Another simple strategy is to take a fake dataset and analyze it to make sure that you can perform the desired analysis. This might take a little extra time, but it could help save more time when you actually have the real dataset in front of you.

Of course, running pilot studies is also very useful. By running one or two pilot participants through the study, you'll be able to identify some of the outstanding issues that you have yet to address in the larger study. It's important to keep the pilot as realistic as possible and to allow enough time to address any issues that arise. Keep in mind that a pilot study is not a substitute for planning ahead. A pilot study is best used to identify the smaller issues that can be addressed fairly quickly before the actual study begins.

11.5 BENCHMARK YOUR PRODUCTS

Usability metrics are relative. There's no absolute standard for what is considered "good usability" and "bad usability." Because of this, it's essential to benchmark the usability of your product. This is constantly done in market research. Marketers are always talking about "moving the needle." Unfortunately, the same is not always true in usability. But we would argue that usability benchmarking is just as important as market-research benchmarking.

Establishing a set of benchmarks isn't as difficult as it may sound. First, you need to determine which metrics you'll be collecting over time. It's a good practice to collect data around three aspects of usability: effectiveness (i.e., task success), efficiency (i.e., time), and satisfaction (i.e., ease of use ratings). Next, you need to determine your strategy for collecting these metrics. This would include how often the data are going to be collected and how the metrics are going to be analyzed and presented. Finally, you need to identify the type of participants to include in your benchmarks (broken up into distinct groups, how many you need, and how they're going to be recruited). Perhaps the most important thing to remember is to be consistent from one benchmark to another. This makes it all the more important to get things right the first time you lay out your benchmarking plans.

Benchmarking doesn't always have to be a special event. You can collect benchmark data (anything that will allow you to compare across more than one study) on a much smaller scale. For example, you could routinely collect SUS data after each usability session, allowing you to easily compare SUS scores across projects and designs. It isn't directly actionable, but at least it gives an indication of whether improvements are being made from one design iteration to the next and how different projects stack up against each other.

Running a competitive usability study will put your data into perspective. What might seem like a high satisfaction score for your product might not be quite as

impressive when compared to the competition. Competitive metrics around key business goals always speak volumes. For example, if your abandonment rates are much higher than your competition, this can be leveraged to acquire budget for future design and usability work.

11.6 EXPLORE YOUR DATA

One of the most valuable things you can do is to explore your data. Roll up your shirt sleeves and dive into the raw data. Run exploratory statistics on your dataset. Look for patterns or trends that are not so obvious. Try slicing and dicing your data in different ways. The keys to exploring your data are to give yourself enough time and not to be afraid to try something new.

When we explore data, especially large datasets, the first thing we do is to make sure we're working with a clean dataset. We check for inconsistent responses and remove outliers. We make sure all the variables are well labeled and organized. After cleaning up the data, the fun begins. We start to create some new variables based on the original data. For example, we might calculate top-2-box and bottom-2-box scores for each self-reported question. We often calculate averages across multiple tasks, such as total number of task successes. We might calculate a ratio to expert performance or categorize time data according to different levels of acceptable completion times. Many new variables could be created. In fact, many of our most valuable metrics have come through data exploration.

You don't always have to be creative. One thing we often do is to run basic descriptive and exploratory statistics (explained in Chapter 2). This is easy to do in statistical packages such as SPSS and even in Excel. By running some of the basic statistics, you'll see the big patterns pretty quickly.

Also, try to visualize your data in different ways. For example, create different types of scatterplots and plot regression lines, and even play with different types of bar charts. Even though you might never be presenting these figures, it helps give you a sense of what's going on.

Go beyond your data. Try to pull in data from other sources that confirm or even conflict with your assertions. More data from several other sources lend credibility to the data you put on the table. It's much easier to commit a multimillion-dollar redesign effort when more than one dataset tell the same story. Think of usability data as just one piece of the puzzle—and the more pieces of the puzzle, the easier it is to fit it all together and get the big picture.

We can't stress enough the value in going through your data firsthand. If you're working with a vendor or business sponsor who ''owns the data,'' ask for the raw data. Canned charts and statistics rarely tell the whole story. They're often fraught with issues. We don't take any summary data at face value; we need to see for ourselves what's going on.

11.7 SPEAK THE LANGUAGE OF BUSINESS

Usability professionals must speak the language of business to have maximum impact. This means not only using the terms and jargon that management understands and identifies with but, more important, adopting their perspective. In the business world, this usually centers on how to decrease costs and/or increase revenue. So if you're asked to present your usability findings to upper management, you should tailor your presentation to focus on how the design effort will result in lower costs or increased revenue. You need to approach usability as an effective means to an end. Convey the perspective that usability is a highly effective way to reach business goals. If you keep your dialogue too academic or overly detailed, what you say probably won't have the impact you're hoping for.

Do whatever you can to tie your metrics to decreased costs or increased sales. This might not apply to every organization but certainly to the vast majority. Take the metrics you collect and calculate how costs and/or revenue are going to change as a result of your design efforts. Sometimes it takes a few assumptions to calculate an ROI, but it's still an important exercise to go through. If you're worried about your assumptions, calculate both a conservative and an aggressive set of assumptions to cover a wider range of possibilities.

Also, make sure the metrics relate to the larger business goals within your organization. If the goal of your project is to reduce phone calls to a call center, then measure task completion rates and task abandonment likelihood. If your product is all about e-commerce sales, then measure abandonment rates during checkout or likelihood to return. By carefully choosing your metrics, you'll have greater impact.

11.8 SHOW YOUR CONFIDENCE

Showing the amount of confidence you have in your results will lead to smarter decisions and help enhance your credibility. Ideally, your confidence in the data should be very high, allowing you to make the right decisions. Unfortunately, this is not always the case. Sometimes you may not have a lot of confidence in your results because of a low sample size or a relatively large amount of variance in the data. By calculating and presenting the confidence intervals, you'll have a much better idea of how much faith or confidence to place in the data. Without confidence intervals, deciding whether some differences are real is pretty much a wild guess, even what may appear to be big differences.

No matter what your data show, show confidence intervals whenever possible. This is especially important for relatively small samples (e.g., less than 20). The mechanics of calculating and presenting confidence intervals is pretty simple. The only thing you need to pay attention to is the type of data you are presenting. Calculating a confidence interval is different if the data are continuous (such as completion time) or binary (such as binary task success). By showing the confidence intervals, you can (hopefully) explain how the results generalize to a larger population.

Showing your confidence goes beyond calculating confidence intervals. We recommend that you calculate p-values to help you decide whether to accept or reject your hypotheses. For example, when comparing average task completion times between two different designs, it's important to determine whether there's a significant difference using a t-test or ANOVA. Without running the appropriate statistics, you just can't really know.

Of course, you shouldn't misrepresent your data or present it in a misleading way. For example, if you're showing task success rates based on a small sample size, it might be better to show the numbers as a frequency (e.g., six out of eight) as compared to a percentage. Also, use the appropriate level of precision for your data. For example, if you're presenting task completion times, and the tasks are taking several minutes, there's no need to present the data to the third decimal position. Even though you can, you shouldn't.

RELATIONSHIP BETWEEN CONFIDENCE INTERVALS AND SIGNIFICANT DIFFERENCES

If the confidence intervals for two means do not overlap, you can assume the two means are significantly different from each other. But if they do overlap, it's still possible that the difference could be significant. A t-test or ANOVA would be needed to check.

11.9 DON'T MISUSE METRICS

Usability metrics have a time and a place. Misusing metrics has the potential of undermining your entire usability program. Misuse might take the form of using metrics where none are needed, presenting too much data at once, measuring too much at once, or over-relying on a single metric.

In some situations it's probably better not to include metrics. If you're just looking for some qualitative feedback at the start of a project, metrics might not be appropriate. Or perhaps the project is going through a series of rapid design iterations. Metrics in these situations might only be a distraction and not add any value. It's important to be clear about when and where metrics serve a purpose. If metrics aren't adding value, don't include them.

It's also possible to present too much usability data at once. Just like packing for a vacation, it's probably wise to include all the data you want to present and then chop it in half. Not all data are equal. Some metrics are much more compelling than others. Resist the urge to show everything. That's why appendices were invented. We try to focus on a few key metrics in any presentation or report. By showing too much data, the most important message is lost.

Don't try to measure everything at once. There are only so many aspects of the user experience that you can quantify at any one time. If a product or business sponsor wants you to capture 100 different metrics, make them justify why each and every metric is essential. It's important to choose a few key metrics for any one

study. The additional time to run the study and perform the analyses may make you think twice about including too many metrics at once.

Don't overrely on a single metric. If you try to get a single metric to represent the entire experience, you're likely to miss something big. For example, if you only collect data on satisfaction, you'll miss everything about the actual interaction. Sometimes satisfaction data might take aspects of the interaction into account, but it often misses a lot as well. We recommend that you try to capture a few different metrics, each tapping into a different aspect of the user experience.

11.10 SIMPLIFY YOUR PRESENTATION

All your hard work comes down to the point where you have to present results. How you choose to communicate your results can make or break a study. There are a few key things you should pay special attention to. First and foremost, you need to understand your audience and their goals.

Often you need to present usability findings to several different types of audiences. For example, you may need to present findings to the project team, consisting of an information architect, design lead, project manager, editor, developer, business sponsor, and prototyper. The project team is most concerned with detailed usability issues and specific design recommendations. Bottom line, they want to know the issues with the design and how to fix them.

TIPS FOR AN EFFECTIVE PRESENTATION OF USABILITY RESULTS

- *Set the stage appropriately.* Depending on your audience, you might need to explain or demo the product, describe the usability testing process, or provide other background information. It all comes down to knowing your audience.

- *Don't belabor procedural details of the test, but make them available.* At a minimum, your audience will usually want to know something about the participants in the test and the tasks they were asked to perform.

- *Lead with positive findings.* Some positive results come out of almost every usability test. Most people like to hear about features of the design that worked well.

- *Use screenshots.* Pictures really do work better than words in most cases. A screenshot that you've annotated with notes about usability issues can be very compelling.

- *Use short video clips.* The days of an elaborate production process to create a highlights videotape are, thankfully, mostly gone. With computer-based video, it's much easier and more compelling to embed short clips directly in the appropriate context of your presentation.

- *Present summary metrics.* Try to come up with one slide that clearly shows the key usability data at a glance. This might be a high-level view of the task completion data, comparisons to objectives, a derived metric representing overall usability, or a usability scorecard.

You also may need to present to the business sponsors or product team. They're concerned about meeting their business goals, participants' reactions to the new design, and how the recommended design changes are going to impact the project timeline and budget. You may present to senior management too. They want to ensure that the design changes will have the desired impact in terms of overall business goals and user experience. When presenting to senior managers, generally limit the metrics and focus instead on the big picture of the user experience by using stories and video clips. Too much detail usually doesn't work.

When presenting results, it's important to keep the message as simple as possible. Avoid jargon, focus on the key message, and keep the data simple and straightforward. Whatever you do, don't just describe the data. It's a surefire way to put your audience to sleep. Develop a story for each main point. Every chart or figure you show in a presentation has a story to it. Sometimes the story is that the task was difficult. Explain why it was difficult, and use metrics, verbatims, and video clips to show why it was difficult, and possibly even highlight design solutions. Paint a high-level picture for your audience. They will want perhaps two or three findings to latch onto. By putting all the pieces of the puzzle together, you can help them move forward in the decision making.