Prototyping

Presentation - User Experience (UX) Design and Management

Luca Leon Happel

2024-11-04 08:00

Presentation website

Please find the interactive version of the presentation here:

https://lucah.tech/prototyping/index.html

Or scan the following QR code:

Figure 1: QR code for the presentation

Or see the source-code on GitHub:

https://github.com/quoteme/prototyping

Presentation notes

Motivation

- large field / may become its own course in the coming years
  –

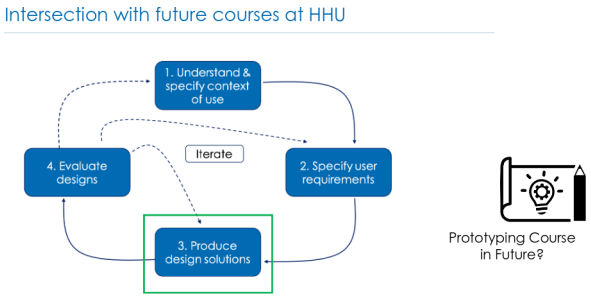


Figure 2: screenshot from `02_UXDM.pdf` from ILIAS

Part of various design paradigms

- UCD (User-Centered Design)
- reason:
  - Reminder: "User-centered design (UCD) describes a process that generates products and services with a high degree of User Experience (UX) by focusing on gaining a deep understanding of users." (Weichert et al. 2019)
  - Prototyping involves users in the design process

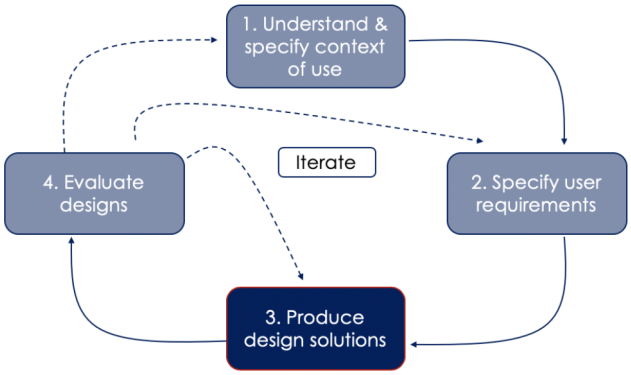
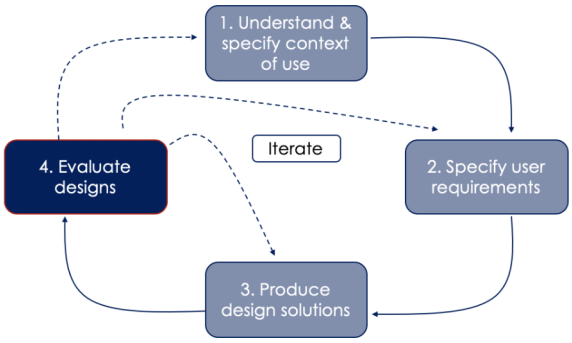


Figure 3: Phase 3/4 in the UCD process

- – reminder:
  - develop various *prototpyes*
  - use prototypes to *communicate* with
    - stakeholders
    - participants
    - decision-makers
    - users
  - develop *common understanding*



- also
  - reminder:
    - use prototypes for *usability testing*
    - *actual users* test the prototypes
    - in *realistic scenarios*

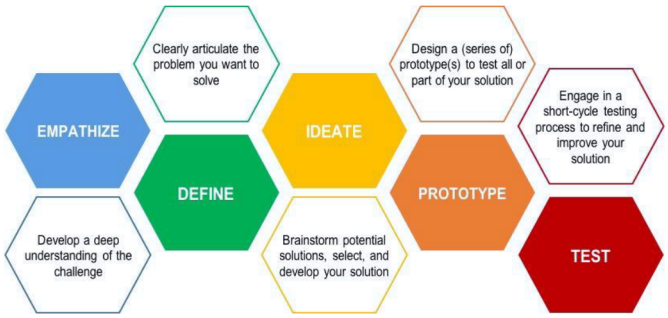


Figure 4: part of design thinking

- – reminder:
    - designed to improve *innovation*
- – reminder:
    - *iterative* development
    - *incremental* development
    - *flexible* development

Takes different forms [1]

- **Proof-of-Principle Prototype**: Verifies key functions
- **Working Prototype**: Near-final functionality
- **Visual Prototype**: Shows size, appearance

[1] Lai, Chun Sing; Locatelli, Giorgio (February 2021). "Valuing the option to prototype: A case study with Generation Integrated Energy Storage" (PDF). Energy. 217: 119290. doi:10.1016/j.energy.2020.119290.

Figure 5: part of agile development

  - **Form Study Prototype**: Emphasizes geometry & less color/material
- **User Experience Prototype**: For user research
  - **Usability Testing**
  - **Realistic Interactions**
  - **Feedback Collection**
  - **Iterative Improvement**
  - **User-Centered Design**
- **Functional Prototype**:
  - Shows function, appearance / not final scale, manufacturability, ...
- **Paper Prototype**:
  - Early UI testing on paper. Confirm and communicate design ideas

Bias

Artifact: The nature of the prototype or product you are evaluating has a huge impact on your findings. The type of interaction will vary tremendously whether it is a paper prototype, functional or semifunctional prototype, or production system.

Figure 6: Artefakte und Bias durch verschiedene Prototypen

Measuring the user experience

well prepared to carry out any UX study involving metrics. In the end, you will likely save time and money and have a greater impact on the product.

3.1 Study Goals

The first decision to make when planning a study is how the data will ultimately be used within the product development life cycle. There are essentially two ways to use UX data: formative and summative.

3.1.1 Formative Usability

When running a formative study, a UX specialist is much like a chef who checks a dish periodically while it's being prepared and makes adjustments to impact the end result positively. The chef might add a little salt, then a few more spices, and finally a dash of chili pepper right before serving. The chef is evaluating, adjusting, and reevaluating periodically. The same is true in formative usability. A UX professional, like a chef, evaluates a product or design periodically while it is being created, identifies shortcomings, makes recommendations, and then repeats the process, until, ideally, the product comes out as close to perfect as possible.

What distinguishes formative usability is both the iterative nature of the testing and when it occurs. The goal is to make improvements in the design prior to release. This means identifying or diagnosing the problems, making and implementing recommendations, and then evaluating again. Formative usability is always done before the design has been finalized. In fact, the earlier the formative evaluation, the more impact the usability evaluations will have on the design.

Here are a few key questions you will be able answer with a formative approach:

• What are the most significant usability issues preventing users from accomplishing their goals or resulting in inefficiencies?

• What aspects of the product work well for the users? What do users find frustrating?

• What are the most common errors or mistakes users are making?

• Are improvements being made from one design iteration to the next?

• What usability issues can you expect to remain after the product is launched?

The most appropriate situation to run a formative usability study is when an obvious opportunity to improve the design presents itself. Ideally, the design process allows for multiple usability evaluations. If there's no opportunity to impact the design, then running a formative test is probably not a good use of time or money. Generally, though, selling the value of formative usability shouldn't be a problem. Most people will see the importance of it. The biggest obstacles tend to be a limited budget or time rather than a failure to see the value.

3.1.2 Summative Usability

Continuing with our cooking metaphor, summative usability is about evaluating the dish after it comes out of the oven. The usability specialist running a summative test is like a food critic who evaluates a few sample dishes at a restaurant or perhaps compares the same meal in multiple restaurants. The goal of summative usability is to evaluate how well a product or piece of functionality meets its objectives. Summative testing can also be about comparing several products to each other. Although formative testing focuses on identifying ways of making improvements, summative testing focuses on evaluating against a set of criteria. Summative usability evaluations answer these questions:

• Did we meet the usability goals of the project?

• What is the overall usability of our product?

• How does our product compare against the competition?

• Have we made improvements from one product release to the next?

Running a successful summative usability test should always involve some follow-up activities. Just seeing the metrics is usually not enough for most organizations. Potential outcomes of a summative usability test might be securing funding to enhance functionality on your product, launching a new project to address some outstanding usability issues, or even benchmarking changes to the user experience against which senior managers will be evaluated. We recommend that follow-up actions be planned along with any summative usability study.

Formative and Summative Usability Testing

The terms *formative* and *summative* were borrowed from the classroom environment, where formative assessment is done on an ongoing basis by a teacher every day in the classroom (think informal observation and "pop quizzes"), while summative assessment is done at the end of some significant period of time (think "final exams"). The earliest application of these terms to usability testing appears to be in a paper presented by Tom Hewett at a conference at the University of York in the United Kingdom (Hewett, 1986). This was also when one of us (Tullis) first met Tom Hewett, mainly because we were the only two Americans at the conference! We've been friends ever since.

3.2 User Goals

When planning a usability study, you need to understand the users and what they are trying to accomplish. For example, are users required to use the product every day as part of their job? Are they likely to use the product only once or just a few times? Are they using it frequently as a source of entertainment? It's critical to understand what matters to the user. Does the user simply want to complete a task or is its efficiency the primary driver? Do users care at all about the design aesthetics of the product? All these questions boil down to measuring two main aspects of the user experience: performance and satisfaction.

3.2.1 Performance

2

Formative effects of prototyping

- reminder:
  - "To find out *improvement areas* prior to release" (`03_UXDM.pdf`)

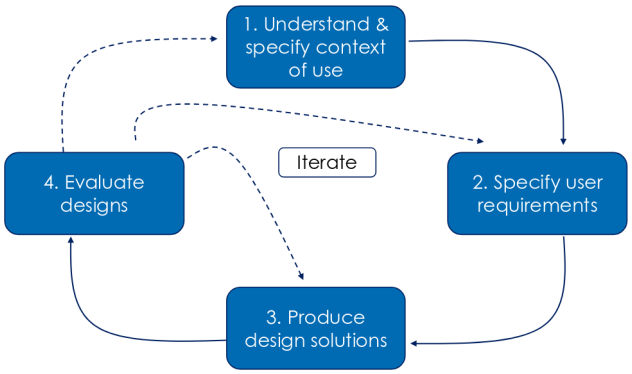


Figure 7: Part 3/4 of the UCD process

*

[2]Thomas Tullis, William Albert, Interactive technologies Elsevier / Morgan Kaufmann, (2013) ISBN:0124157815,9780124157811,0124157920,9780124157927

- Key Questions & Answers:
  - **What usability issues block user goals?**
    Discover issues through prototype testing
  - **What aspects work well or frustrate?**
    Observe user interactions with prototypes
  - **What common errors are users making?**
    Identify errors in prototype use
  - **What issues may remain post-launch?**
    Predict challenges from prototype feedback
  - **How can prototypes enhance understanding?**
    Use prototypes to align with users
  - **What feedback can improve design?**
    Iterate based on prototype insights

Summative effects of prototyping

- reminder:
  - "To determine the *quality of UX* against a set of criteria" (`03_UXDM.pdf`)

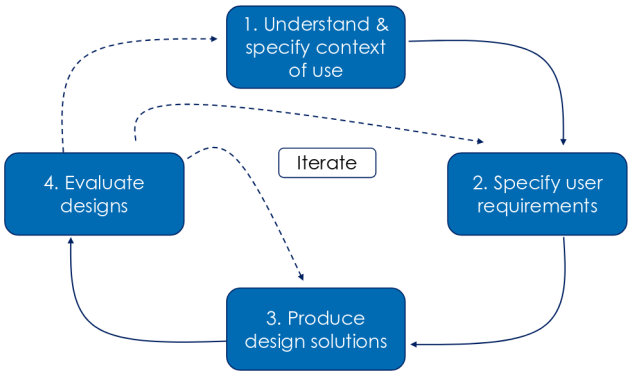


Figure 8: Part 4/4 of the UCD process

*

- Key Questions & Answers:
  - **Did we meet UX/usability goals?**
    Verify goals using final prototype tests
  - **What is the overall UX quality?**
    Evaluate user experience with polished prototype
  - **How does our product compare?**
    Benchmark prototype against competitor models
  - **Have we improved since last release?**
    Compare current prototype to prior versions

Applications of prototypes

- comparing designs

much easier, more efficient, or more entertaining than expected, you know you are onto something.

Another set of self-reported metrics relates to future use. For example, you might ask questions related to likelihood to purchase, recommend to a friend, or use in the future. The Net Promoter Score is a widely used metric to measure likelihood of future use. Another interesting set of metrics relates to subconscious reactions that users may be having. For example, if you want to make sure your product is engaging, you can look at physiological metrics. Changes in pupil diameter can be used to gauge the level of arousal or, if you're trying to eliminate stress as much as possible, you can measure heart rate or skin conductance.

3.3.9 Evaluating the Impact of Subtle Changes

Not all design changes have an obvious impact on user behavior. Some design changes are much more subtle, and their impact on user behavior is less clear. Small trends, given enough users, can have huge implications for a large population of users. The subtle changes may involve different aspects of the visual design, such as font choice and size, placement, visual contrast, color, and image choice. Nonvisual design elements, such as subtle changes to content or terminology, can also have an impact on the user experience.

Perhaps the best way to measure the impact of subtle design changes is through live-site metrics from A/B tests. A/B testing involves comparing a control design against an alternative design. For websites, this usually involves diverting a (usually a small) portion of web traffic to an alternative design and comparing metrics such as traffic or purchases to a control design. An online usability study with a large population can also be very useful. If you don't have access to the technology to run A/B tests or online studies, we recommend using e-mail and online surveys to get feedback from as many representative participants as you can.

3.3.10 Comparing Alternative Designs

One of the most common types of usability studies involves comparing more than one design alternative. Typically, these types of studies take place early in the design process, before any one design has been fully developed. (We often refer to these as "design bakeoffs.") Different design teams put together semifunctional prototypes, and we evaluate each design using a predefined set of metrics. Setting up these studies can be a little tricky. Because the designs are often similar, there is a high likelihood of a learning effect from one design to another. Asking the same participant to perform the same task with all designs usually does not yield reliable results, even when counterbalancing design and task order.

There are two solutions to this problem. You can set up the study as purely between subjects, whereby each participant only uses one design, which provides a clean set of data but requires significantly more participants. Alternatively, you can ask participants to perform the tasks using one primary design (counterbalancing the designs) and then show the other design alternatives and ask for their preference. This way you can get feedback about all the designs from each participant.

The most appropriate metrics to use when comparing multiple designs may be issue-based metrics. Comparing the frequency of high-, medium-, and low-severity issues across different designs will help shed light on which design or designs are more usable. Ideally, one design ends up with fewer issues overall and fewer high-severity issues. Performance metrics such as task success and task times can be useful, but because sample sizes are typically small, these data tend to be of limited value. A couple of self-reported metrics are particularly relevant. One is asking each participant to choose which prototype he would most like to use in the future (as a forced choice comparison). Also, asking each participant to rate each prototype along dimensions, such as ease of use and visual appeal, can be insightful.

3.4 Evaluation Methods

One of the great features of collecting UX metrics is that you're not restricted to a certain type of evaluation method (e.g., lab test, online test). Metrics can be collected using almost any kind of evaluation method. This may be surprising because there is a common misperception that metrics can only be collected through large-scale online studies. As you will see, this is simply not the case. Choosing an evaluation method to collect metrics boils down to how many participants are needed and what metrics you're going to use.

3.4.1 Traditional (Moderated) Usability Tests

The most common usability method is a lab test that utilizes a relatively small number of participants (typically 5 to 10). The lab test involves a one-on-one session between a moderator (usability specialist) and a test participant. The moderator asks questions of the participants and gives them a set of tasks to perform on the product in question. The test participant is likely to be thinking aloud as she performs the various tasks. The moderator records the participant's behavior and responses to questions. Lab tests are used most often in formative studies where the goal is to make iterative design improvements. The most important metrics to collect are about issues, including issue frequency, type, and severity. Also, collecting performance data such as task success, errors, and efficiency may also be helpful.

Self-reported metrics can also be collected by having participants answer questions regarding each task or at the conclusion of the study. However, we recommend that you approach performance data and self-reported data very carefully because it's easy to overgeneralize the results to a larger population without an adequate sample size. In fact, we typically only report the frequency of successful tasks or errors. We hesitate even to state the data as a percentage for fear that someone (who is less familiar with usability data or methods) will overgeneralize the data.

Usability tests are not always run with a small number of participants. In some situations, such as comparison tests, you might want to spend some extra time and money by running a larger group of participants (perhaps 10–50 users). The main advantage of running a test with more participants is that as your sample size increases, so does your confidence in your data. Also, this will afford you the ability to collect a wider range of data. In fact, all performance, self-reported, and physiological metrics are fair game. But there are a few metrics that you should be cautious about. For example, inferring website traffic patterns from usability-lab data is probably not very reliable, nor is

Figure 9: Albert and Tullis - 3.3.10

–

- online surveys

3.4.3 Online Surveys

Many UX researchers think of online surveys strictly for collecting data about preferences and attitudes, and firmly in the camp of market researchers. This is no longer the case. For example, many online survey tools allow you to include images, such as a prototype design, within the body of the survey. Including images within a survey will allow you to collect feedback on visual appeal, page layout, perceived ease of use, and likelihood to use, to name just a few metrics. We have found online surveys to be a quick and easy way to compare different types of visual designs, measure satisfaction with different web pages, and even preferences for various types of navigation schemes. As long as you don't require your participants to interact with the product directly, an online survey may suit your needs.

The main drawback of online surveys is that the data received from each participant are somewhat limited, but that may be offset by the larger number of participants. So, depending on your goals, an online survey tool may be a viable option.

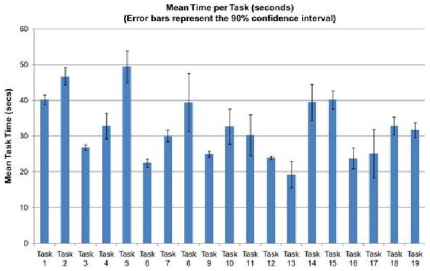
Figure 10: Albert and Tullis - 3.4.3

–

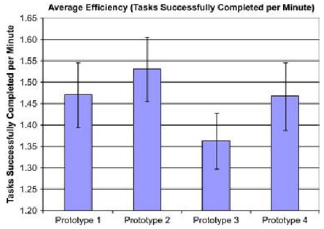
– 

[3]

- performance metrics
  - user efficiency



∗ **FIGURE 4.4** Mean time on task, in seconds, for 19 tasks. Error bars represent a 90% confidence interval. These data are from an online study of a prototype website. [4]



∗ **FIGURE 4.10** Average number of tasks completed successfully per minute in an online study of four different prototypes of navigation for a website. Over 200 participants attempted 20 tasks with each prototype. Participants using Prototype 2 were significantly more efficient (i.e., completed more tasks per minute) than those using Prototype 3. [5]

- testing
  - user success in A/B testing

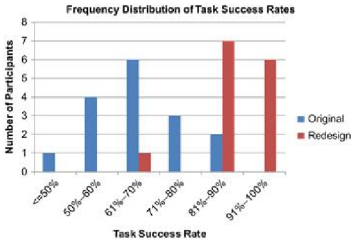
∗ Task success for different groups of participants is also used when each group is given a different design to work with. For example, participants in a usability study might be assigned randomly to use either Version A or Version B of a prototype website. A key comparison will be the average task success rate for participants using Version A vs those using Version B. [6]

[3] Thomas Tullis, William Albert, Interactive technologies Elsevier / Morgan Kaufmann, (2013) ISBN:0124157815,9780124157811,0124157920,9780124157927 - chapter 3

[4] Thomas Tullis, William Albert, Interactive technologies Elsevier / Morgan Kaufmann, (2013) ISBN:0124157815,9780124157811,0124157920,9780124157927 - chapter 4

[5] Thomas Tullis, William Albert, Interactive technologies Elsevier / Morgan Kaufmann, (2013) ISBN:0124157815,9780124157811,0124157920,9780124157927 - chapter 4

[6] Thomas Tullis, William Albert, Interactive technologies Elsevier / Morgan Kaufmann, (2013) ISBN:0124157815,9780124157811,0124157920,9780124157927 - chapter 4



∗ [7]

Märchenstunde dazu

- Prototype of a high-voltage detector
- Phase 2/4 UCD
  - Discussion with developers

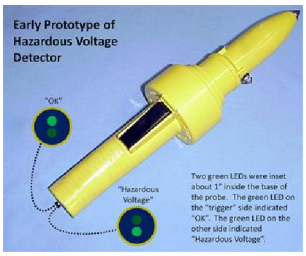


Figure 11: LEDs

    - ∗
    - ∗ If wrong LED is on, user could die
  - Developers did *not want to change* LEDs
- Phase 3/4 UCD
  - Prototyp has been build
    - ∗ **Simulates** 20% probability, measured object is high voltage
- Phase 4/4 UCD
  - 100 Participants
  - One misinterpretation

Resultat Developers got convinced to change LEDs

---

[7] Thomas Tullis, William Albert, Interactive technologies Elsevier / Morgan Kaufmann, (2013) ISBN:0124157815,9780124157811,0124157920,9780124157927 - chapter 4