Transforming Publishing User Research: User Testing

<table>
<thead>
<tr>
<th>Document Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>Version 1.0</td>
</tr>
<tr>
<td>Date Issued</td>
<td>July 2018</td>
</tr>
<tr>
<td>Author</td>
<td>John Mowbray</td>
</tr>
<tr>
<td>Comments to</td>
<td><a href="mailto:nss.isdtransformingpublishing@nhs.net">nss.isdtransformingpublishing@nhs.net</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Comment</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1.0</td>
<td>July 2018</td>
<td>First draft</td>
<td>John Mowbray</td>
</tr>
<tr>
<td>Version 1.1</td>
<td>July 2018</td>
<td>Edited following comments on first draft from Anna Price</td>
<td>John Mowbray</td>
</tr>
</tbody>
</table>
Contents

1. Introduction............................................................................................................................................. 3
2. What is a usability test................................................................................................................................. 3
  2.1.1 Planning stage ................................................................................................................................. 3
  2.1.2 Recruiting participants .................................................................................................................... 4
  2.1.3 Test design ....................................................................................................................................... 5
  2.1.4 Test implementation ........................................................................................................................ 7
  2.1.5 Analysis and outputs ......................................................................................................................... 8
1. Introduction

This document provides ISD analysts with advice on conducting user research when transforming publications into a new interactive format. It covers one of the two main types of research to be conducted during this process: User Testing (UT). Guidance on conducting initial user engagement can be found in the “Initial User Engagement” document here.

It should be noted that the guide is not exhaustive, nor is it prescriptive. Indeed, to adapt to a range of software development products (e.g. a data trends webpage, or a data explorer/dashboard), UT is a necessarily diverse practice. As such, this report covers some of the key principles that should be considered when conducting this type of user research. However, it is strongly encouraged that any teams undertaking user research also conduct additional background research, to develop user tests that gather meaningful data about their specific products. Links to supporting documentation and resources that are useful for further reading are provided throughout this report.

2. What is a usability test

UT (also known as formative testing) is undertaken during the development process. First and foremost, it helps to diagnose and fix problems with prototypes of a software product (e.g. a data explorer) to determine whether the product is usable. It is also a useful opportunity to determine whether the new product is efficient, effective, and if it provides satisfaction to the user. This latter point is especially important – i.e. that the new product should provide more value than the old one. It may be useful to conduct several user tests to test the product as it develops.

The following quotation is useful to bear in mind when questioning the needs and/or benefits of usability tests:

“From the moment you know enough to talk about a product—any product, whether it’s hardware, software, a video game, a training guide, or a website—you know too much to be able to tell if the product would be usable for a person who doesn’t know what you know” (Nielsen, 1993).

It is important to bear in mind that the main purpose of a user test is not:

- To learn about user journeys
- To gain content ideas
- To provide a demonstration of the prototype
- To seek validation for your work

These may be indirect outcomes of the test; however, its main purpose is learning how to refine the prototype to make it more user friendly. Therefore, it must be approached with a view to inviting critique, and with an objective and open mind.

2.1.1 Planning stage

Meetings to plan UTs should take place early in the development process. At a minimum they should cover the following topics:
1. **The definition of objectives**

The overarching goal of UT is to determine the extent to which your product is usable. However, it is important to define more specific objectives that will help you to achieve this goal. To this end, it is necessary to identify functions or aspects of the software that you would like to test during the first iteration¹. For example, if testing a prototype of a data explorer, you may wish to focus on the ease with which users can operate the system to a) create data visualisations (i.e. graphs and charts) using drop down menus; b) use visualisations to source specific pieces of information; and c) use advanced functionality e.g. downloading .png images of visualisations.

2. **The designation of a project lead**

This individual (or these individuals) should be tasked with the responsibility of managing the UT process from the analytical team’s perspective, often working in conjunction with/receiving support from the Transforming Publications team. This is an important role, as conducting user research involves multiple tasks which can be time-consuming and challenging. The intensity of work will ebb and flow throughout, and busy periods can be difficult to predict. This is because a significant proportion of the work requires the input of individuals from inside and outside of the organisation.

3. **Agreement on general principles and next steps**

It may not be a productive use of time to create a thorough plan for the implementation UT. Conducting research can be a complex and largely iterative exercise, where it is necessary to react to various situational factors. Instead, it would be useful to agree general principles (e.g. relating to the documentation of activities, providing periodic feedback to the team, support networks, and communication channels), and agreeing on initial tasks that need to be undertaken. It is particularly important that the project lead has access to support, as many activities will require input and feedback from the team.

2.1.2 Recruiting participants

Participant recruitment is an important part of UT, and should start as early in the overall process as possible. For example, a good time to begin recruiting participants is at the outset of the research design phase (see Section 2.1.3). This is because recruitment is often a lengthy and staggered process, involving a significant degree of backwards and forwards communication with potential participants. It also involves various administrative exercises, such as room-booking, journey planning, and completing expense claims (e.g. for travel costs). Therefore, it is more efficient to undertake these activities in tandem with others at the preparation phase, in order to establish a list of willing participants.

Most stages of UT only require a small number of participants. Studies indicate that – in a specific test but not a test of the whole product - 85% of usability problems will be identified by five participants (Barnum, 2011). Despite this, it is still necessary to carefully consider the amount and type of the participants that would be needed to provide a valuable critique of your product’s usability. Figure 1 contains a case study of the considerations that could be applied to UT.

¹ Note that it may be necessary to conduct several tests as the product develops. These could be used to test newly developed features, or different parts of the platform/webpage.
It should be emphasised that the purpose of the case study in Figure 1 is to highlight the need for a cogent strategy and reasoning when preparing a research project. However, the actual decisions taken may differ from the case outlined above, depending on the context of the test. For example, you may be undertaking a second iteration of a user test (i.e. having completed a first iteration and made subsequent changes to the product) and be keen to target a larger sample of non-technical users. This decision may be taken because it has been determined from the previous test that non-technical users in particular face barriers to using the product. Therefore, although there are basic guiding principles to participant recruitment, different cases provide unique circumstances.

It should also be noted that, although previous research suggests that in the majority of cases only a small number of participants is needed to uncover most usability issues, this may not apply in every single case. The following question should be considered when testing has finished:

Towards the end of the testing phase, were significant new faults or issues with the system still being found?

If the answer is yes, then it is advisable to run more tests. Some further development work could be undertaken beforehand (and whilst new participants are being recruited), to address some of the usability issues that have already been identified.

2.1.3 Test design

When designing UT, there are three key documents that should be prepared. These are as following:

---

**Figure 1: sample considerations case study**

**Context:** a data explorer prototype is being tested by the Mental Health publications team, to be included as part of their transformed publication format. They establish the following parameters for recruiting a sample of participants:

1. To target a total of 10 participants. This will ensure that at least 85% of usability problems will be identified, and that certainly the most pressing issues will be discovered during the tests.
2. Up to half of the sample should be recruited from within ISD, to provide an insight into how users who are familiar with data analysis experience using the product. Preferably, they should not have previous knowledge of the prototype or its functionality.
3. At least half of the sample should be from a non-ISD background. This would provide an insight into how users who are non-technical, or who are less familiar with data analysis, experience using the product. The results would provide an interesting comparison with the sample group outlined in point 2.
A test script for the user

This should contain a list of tasks for the participant to undertake. This document is known as a ‘schedule’ or a ‘script’, and is used by the moderator to guide the user through the test. The tasks in the session should:

- Be clear and understandable
- Flow in a natural and logical manner
- Be sufficient to address the research objectives

It should also be remembered that – as much as possible - the testing session should replicate the user’s natural experience. To this end, the following basic principles should be applied to script design:

Tasks should reflect typical user journeys (identified via initial user engagement sessions). For example, users may come to your site to find a specific type of information. In this case, a suitable task could be “for the year 2005/06, please find the number of inpatients in the Greater Glasgow & Clyde health board”

Tasks should not be too prescriptive i.e. they should not tell the user what do in a step by step manner. It may be that the task can be completed: (a) in one way only, (b) in various efficient ways, or (c) in one efficient way and various inefficient ways. It is important to observe how users tackle the activity, and to note any barriers or issues that they face.

A test script for the facilitator

This should be used by the facilitator as a guide during the test. It will follow the same structure as the script for users, but will contain space for:

- Recording completion rates for tasks
- Handwritten observation notes about task completion

The script should also contain examples of follow-up questions. For example, after each activity (or after a section of activities) it may be useful to ask the participants for their thoughts on completing the task. Questions can also be asked at the end of the test, to find out how they found the entire experience of using the system, or how it compares with the current/traditional system that it is replacing. The amount of questions should be kept to a minimum – especially those in between activities.

Closed-ended questions (i.e. that elicit yes/no answers) should be avoided. Instead, questions should be tailored to prompt feedback. For example, instead of asking “Did you enjoy using the explorer?”, you could ask “Now that you have completed the test, what are your thoughts on using the explorer?” When asking follow-up questions, be ready to ask probing questions (e.g. that’s interesting, could you tell me a bit more about that?) if necessary.

---

2 These objectives may evolve during the planning phase or after initial testing has been completed. However, the latest objectives should always be used to keep the research on track.
A quantitative feedback form

In addition to handwritten notes about the tests, it is also good practice to gather some objective feedback from the participants during the tests. This information will be particularly useful when it comes to analysing, interpreting, and reporting the test findings (see Section 2.1.5). A simple ease of use rating scale can be a useful means of gathering this quantitative data. On a scale of 1-7 (very difficult to very easy), it allows users to provide a direct score for each activity.

At the end of the test it may also be appropriate to gather some objective feedback about the users’ overall experience of engaging with the system. The System Usability Scale is a quick and effective means of gathering this data. It rates 10 different features relating to the software, such as its ease of use and how much the participant enjoyed using the system.

2.1.4 Test implementation

To avoid participant fatigue, a user test should last no longer than one hour. Ideally, there will be two people facilitating the session. This will ease the burden of note taking, and will ensure that two differing perspectives are captured, adding more weight and validity to the findings. Having two people involved is also useful when it comes to practical arrangements – e.g. setting up technology and preparing materials. However, be mindful that participants may feel under pressure when being observed by two people. To this end, a few minutes of informal conversation before the test gets underway can help to break the ice. Then, when turning attention to the task at hand, it is good practice to: a) explain in simple terms the purpose and background of the research; and, b) give clear details about the format of the test. Assume that the user knows nothing about the project, and nothing about the nature/purpose of UT. Also, remember to state the following mantra:

“The purpose of the session is to test the software, NOT you”.

In addition to the above, assure the participants that everything they say and do during the test is confidential, and that their identities will be anonymised during the analysis and reporting stages of the process. If you intend to record the session using either audio or visual technology, be sure to ask participants for permission, and make it clear that they have a right to refuse. While such recordings may be useful when it comes to maximising the findings from the tests (i.e. making sure that nothing is missed in handwritten notes), keep in mind that they may distract or put unnecessary pressure on participants – in which case, they may actually have a deleterious impact on test outcomes.

During the session itself, it is probable that the participant will ask for help. In these cases, it is crucial not to intervene. For example, if a participant asks “how do I do this?”, you could respond by saying “what is your instinct?”. In real life, your users will not have the benefit of a demonstrator to assist them; therefore, by providing assistance, you will not be learning about your product’s actual usability. If participants are inclined to give up a task quickly, perhaps encourage them to keep exploring for another minute or so to see if they can complete it. This will allow you to observe how users engage in trial and error with your prototype, and to understand the scale of the functionality issue.
Another useful means of capturing data during the tests (in addition to notes, recordings, and quantitative rating scales), is to encourage participants to think aloud when engaging in test activities. This is a fairly standard approach for UT, where users talk through the process of completing the activities. There are many resources online that describe and critique the think aloud method (for example, scholarly resources, and those resources geared towards practitioners).

At the end of the session the participants should be informed of what happens next, and how they will be kept involved in the project. It is good practice to make participants feel valued to be kept informed of progress. It may also be an opportunity to determine whether they would take part in future usability tests or other forms of user engagement. These details should be recorded for future reference.

2.1.5 Analysis and outputs

After the usability tests it is important to analyse the data you have gathered to determine next steps in terms of designing your new product. Firstly, you must transcribe any audio recordings and compile observational notes. Following this, you must code (or categorise) the data. At this stage, it may be useful to divide the results in accordance with different tasks carried out during the test. For example, when testing a data explorer, you may categorise the results as following:

- Reading the introductory landing page
- Creating data visuals and finding information
- Using advanced functionality
- Miscellaneous category

Having done this, it will be necessary to read notes and transcripts thoroughly, and allocate important quotes, insights, or other findings to each of these categories. You may then further code this data, by developing themes e.g. problems identified, usability suggestions provided by participants, content requests etc. Be sure to quantify the comments that underpin themes, to give some indication about the importance of the different issues encountered. For example, count how many times specific problems were identified by different users, or how often a specific design suggestion was being made.

If quantitative data were gathered during the test, this can be inputted into a simple spreadsheet for analysis. These results can then be used to triangulate the qualitative data (i.e. observation notes/transcripts). For example, you may have observed issues with each of the different tasks allocated in the test. If the participants also rated the ease of task completion on a rating scale of 1-5, this information will indicate the severity of the issues they faced. For example, one task may have been allocated an average ease of use score of 4 out of 5, whilst another received a 2 out of 5. In this case, it may be necessary to focus more upon the problems relating to the latter task.

In addition to the above, it is important to generate a report on the overall research process. This should include a description of the background, methodology (including a sample description), findings, and practical recommendations for future design work. It may also be useful to have a separate summary document that states only the key findings and recommendations. Such documentation will be invaluable as an audit trail, which shows
exactly how the usability tests were carried out, and how they influenced the design process. It is also a record that can then be used to demonstrate how the product provides value to users, and can also potentially be used by other teams in the organisation embarking upon similar development work.

Resources

There are some absolutely fantastic resources for usability testing online. The following four should be a good place to start.

This is a guide to help someone managing a usability test.

Usability.gov has some excellent resources on designing and conducting UTs

A very useful and more detailed guide to usability testing

This is a document on user testing for those using an Agile approach to project management