Usability testing

What is usability?

Usability is the study of how people interact with the objects around them, whether that’s a website, an application, a toaster or a door. Every man-made item we own was designed by someone and each can be evaluated in terms of its usability.

There are many definitions of what the term ‘usability’ means.

The International Standards Organisation defines it as “the extent to which a product can be used by specified users to achieve goals with effectiveness, efficiency, and satisfaction in a specified context of use.”¹ The Usability Professionals’ Association refers to it as “the degree to which something - software, hardware or anything else - is easy to use and a good fit for the people who use it.”²

There are countless other definitions, but they all have three common aspects:

• There’s a user.
• The user is trying to achieve something.
• The user is interacting with an object, product, service or system.

An example: the Amenities building automatic door

Last year, the University of Nottingham extended Jubilee campus and added a number of new buildings. These look funky and modern, but whoever designed them seems to have forgotten one key detail: the doors aren’t easy to use.

¹ International Standards Organisation - ISO 9241-11
In the Amenities building, you’ll likely to see the following situation. Someone will try to exit the building through the doors to Triumph Road, but they can’t. They walk up to the door, expecting it to open, but nothing happens.

What they’ve failed to notice is the pillar on the right hand side of the door. On top of this, there’s a tiny button which once pressed, opens the door.

There are a number of issues with this. First of all, the labelling on the pillar is too small and set in low contrast colours that make it difficult to pick out. There’s plenty of space that could have been used to make the text larger and it would have been trivial to choose a better colour. Imagine that your vision is less than perfect; would you be able to read that text?

You can approach this exit from one of three directions: left, right or where I’ve taken the photo from. Unless you’re coming from the right, you have to go out of your way to press the button on the pillar. If you’re coming from the left, you have to walk across the door, press the button and then go back towards the middle. This doesn’t seem too bad for able-bodied people, but anyone with mobility problems might find this frustrating, especially after repeat use.

The labelling that is provided is also misleading. When you see the word ‘automatic’, it essentially means ‘no further action required’. When coupled with the word ‘door’, it creates a strong mental model in our heads which has been reinforced by all of the other things we’ve seen labelled “automatic door” before. To open the door, you walk up to it and it opens. This is an interaction which is consistent across almost all doors labelled in this way, so when it doesn’t do what we expect, it’s surprising and confusing.

Doors that don’t work the way we expect them to aren’t the end of the world, but they make our lives that little bit more frustrating. For people with visual impairments or who have problems with mobility, it can be even more so.
Why is usability important?

While it might not seem like the most obvious thing to consider when writing software or creating any other product, the consequences of ignoring usability can be very high and the rewards for embracing it can be equally great.

Life and death

There are many situations where people control objects or systems that, if they failed, would cause huge losses of human life, environmental damage or other significant consequences.

For example, if a pilot crashes a plane, hundreds of people could lose their lives. If the controllers of a nuclear power station accidentally cause a meltdown, it can expose hundreds of thousands of people to radiation. If a submarine operator accidentally launches a nuclear warhead, then it could lead to an even greater disaster.

Although the cost of making mistakes in these examples is extremely high, in almost all cases, these systems still need human input to operate. While this is the case, the upmost must be done to ensure that as few errors are made as possible and when they are made, they are easy to recover from.

Money

One of the main incentives for companies to invest in usability is money. Almost all companies derive revenue from activities which involve human interaction, whether that be with customers, other businesses or their own employees. Making these interactions as easy as possible can dramatically improve a company’s fortunes by increasing sales, productivity and return business.

Consider the online retailer Amazon. Their entire business revolves around their website, which customers interact with. If it was difficult and frustrating to use, fewer people would buy goods from them and they would quickly go out of business.

The same principles apply to internal systems in companies. If a business provided employees with an email program that isn’t usable, then it would considerably reduce their productivity. People will go to great lengths to avoid using difficult systems and if they’re forced to use them, they’ll do so as little as possible, waste time doing so and require more support.³

On the other hand, companies can profit from good usability. By concentrating on providing an excellent user experience, businesses such as Apple and Google have succeeded where their competitors have failed.

**How can you determine usability?**

Usability doesn’t initially seem like the easiest thing to measure or determine. It’s a relatively abstract concept and not as straightforward as counting lines of code or unit test coverage. However, there are a number of methods which can be incorporated to determine usability.

**Expert review**

Like many things, the most straightforward way to find out if something is usable or not is to ask an expert. Typically this would involve a professional usability analyst examining the product or system in question and producing a report which highlights the good and bad points of it, plus recommendations. However, although this can be the quickest way to get feedback, its obvious limitation is that it’s only one person’s view. Despite their knowledge, they are unlikely to be a representative user of the system, so may miss some of the potential problems in it.

**Interviews, questionnaires and focus groups**

You can also determine how usable people find a system or object by simply asking them. This is useful for finding out the overall level of satisfaction for something and general feelings towards it, but this information isn’t that reliable. This is because we can’t view our interactions with things in an objective way and our memory is unable to remember events as they actually occurred. Jakob Nielsen, a famous usability analyst, concludes, “self-reported claims are unreliable, as are user speculations about future behaviour.”

**Log files and analytics**

One method which is objective is to review user activity log files or website analytics. This shows clearly how people are interacting with the system or object in question, but it doesn’t show why. It may be able to chart a user’s path through a system, but not the reasons why they chose that path. For this reason, it’s often a useful companion to other usability work, but can’t be used on its own.

**Usability testing**

Usability testing is the last and generally most useful way of determining how usable something is. It involves putting participants in situations similar to real life and getting them to perform typical tasks associated with the object or system on test. It is usually the most expensive type of usability analysis, but produces the best results.

---

What is usability testing?

Put simply, usability testing is asking someone to use something and watching them. If software is under test, a participant will usually sit at a computer while a moderator instructs them from beside or behind them, occasionally stopping to ask questions. Using these tests, the design and development team can immediately see if people are using the system or product as intended.6

Participant numbers

Although you’re always likely to find more usability errors the more people you test with, it’s generally accepted that you don’t need more than five participants to find most of them.7 Research by Jakob Nielsen has shown that you can typically find 75% of usability problems with only four users.8

Suitability of participants

Ideally, participants should be as close to the actual users of a product or system as possible, but much of the time, this isn’t possible. However, in most cases, this doesn’t matter because “the importance of recruiting representative users is overrated.”9 When testing software, unless the target demographic is very narrow or domain-specific knowledge is required, most people will be suitable to test with.

Equipment

You can do usability testing with no additional equipment, but it’s useful to record participants so that you don’t miss anything and the test can be analysed in greater detail later. If you’re conducting many tests in a day, you may lose concentration and miss an important insight into the participant’s interaction with your system or product.

To record usability tests conducted on a computer, you need software to record the screen, a camera and software to capture the participant’s face and reactions, plus a microphone to record what is being said.

Professional usability agencies normally use the software package Morae, which retails at £120-£900.10 This allows testers to capture the test, fellow observers to see it in real-time and manage test data. Most professionals also use eye tracking hardware, such as a screen with a built in camera. These produce ‘heat maps’ which show where the user was

7 Tullis and Albert - “Measuring the User Experience” - Morgan Kaufmann, 2008, pages 118-121
8 Nielsen, Jakob - “Why You Only Need to Test with 5 Users” - http://www.useit.com/alertbox/20000319.html
9 Krug, Steve - “Don’t Make Me Think!” - New Riders, 2006, pages 135, 139-140
looking during the test. However, eye tracking hardware and software is expensive, with a setup usually costing several thousand pounds.

It is possible to record usability tests for relatively low cost though. Silverback, by Clearleft, can record the screen, the participant and microphone audio, but only costs £30.\textsuperscript{11} It can then produce a composite video of these three inputs. While it requires an Apple computer, it is much cheaper than alternative products.

**Remote usability testing**

Although most usability testing is done where the participant is in the same room as the moderator, testing can be done online, where both are in different locations. Average broadband connection speeds now allow the participant’s screen to be captured over the Internet with relative stability, although making sure that the participant has the correct hardware and software required can be a logistical barrier to this kind of testing.

**Metrics**

As in many fields, usability has its own set of metrics that can be measured. These include performance, issues-based, self-reported, behavioural, physiological and comparative metrics.\textsuperscript{12} In conjunction with qualitative data, these can help to understand how usable or not a particular product or system is.

**Analysing and using test results**

Once a test is completed, you end up with a large amount of raw data in the form of pre-test comments, post-test comments and recorded video/audio. Qualitative and quantitative data must then be extracted from these by observing participant behaviour and comments, and measuring metrics like success rates and time-on-task.

After data has been analysed and a list of problems and recommendations has been written, the test results need to be communicated to designers, developers and anyone else that needs to know.\textsuperscript{13} This can be done verbally, in writing or visually through a presentation or video.

\textsuperscript{11} Clearleft Silverback - \url{http://silverbackapp.com}

\textsuperscript{12} Tullis and Albert - *Measuring the User Experience* - Morgan Kaufmann, 2008

How does usability testing fit into software quality and assurance?

Usability testing should be considered to be as important as any other quality assurance method. After all, if someone can't use your software, then it's useless.

When to test

Usability testing should be conducted as often as your time and financial budget allows. Tests should also be run as early as possible, as changes to systems are more costly the further they are into development. Steve Krug notes, “testing one user early in the project is better than testing 50 near the end”.\(^\text{14}\)

Testing can be done at many stages of a project. You might consider testing similar products before you design your own, testing prototypes, testing after each development milestone, when you’re doing beta testing and after the product has shipped. Whenever you test, the more you do, the better.

Usability testing within iterative development

Although usability testing is suitable for use with any software development methodology, it can work particularly well with the iterative development model.\(^\text{15}\) As testing usually results in a list of recommended changes, it works best with a flexible development model that allows for multiple alterations.

Summary

Usability examines the things we create in the context of human behaviour. Usability testing provides us with the best way to analyse how objects, products, systems and services perform when people interact with them. Although professional testing can be expensive, it can also be done inexpensively. Usability testing fits into quality assurance alongside other testing methods and is just as important.

\(^\text{14}\) Krug, Steve - “Don't Make Me Think!” - New Riders, 2006, page 134