

## ***2.3.7 Usability tests***

# What do we mean by usability test?

- A usability test is a one-on-one session where a representative user interacts with a prototype design with a researcher facilitating and taking notes
- Information you can get out includes:
  - Whether the **proposition is clear and of real value** to the user
  - Whether there are **usability issues that inhibit successful** use of the website
  - Whether key tasks can be quickly and readily completed by users
- A usability test can be **qualitative**, through open-ended interview questions, or **quantitative** by measuring task success and time rate

# Quantitative testing

- A quantitative test can help inform the final stages of design by **verifying the performance of the site** along the key measures
  - Task completion rate and time
  - Satisfaction rating
- The researcher sets a few **standards tasks, times the participant and writes a survey to capture results** at the end
- When it identifies small problems, quant testing can identify where tweaks need to be made to the design
- But when it identifies bigger problems, quant testing typically doesn't provide enough guidance for a redesign

# Qualitative testing

- Run **more like an open-ended interview**, qualitative testing lets the participants explore the prototype, exploring their own motivations & interests as well as your test objectives
- Qual testing is quite strong at **discovering a more holistic understanding of your user's mental model, perceptions and motivations**
- But you need to **be careful with interpreting** what your users say
- And it is often **hard to get an overall picture** of how well the current design performs

# How to do a usability test

- Mix quantitative and qualitative objectives to suit your evaluation objectives
  - Regardless of what you want to learn and where you are in the design process, it is almost always useful to include both qual and quant methods
- Set up the usability test
  - Users
  - Client observations
  - Write an interview script/ discussion guide
  - Write a brief survey for your quantitative objectives

# Users – who and how many?

- Should be as **representative** as possible
- For an intranet, it's easy to find appropriate users!
- For an Internet site, need to have a **sample of users with similar demographic distribution** to the intended user population
- Normally you'll get those from the **specialist market research recruiter** your usability consultants work with
- Jakob says you only need to test with 5 users, but not everyone agrees!
- A normal test is conducted with 5-10 users doing the same things

# Observing sessions

It's important that stakeholders turn up to observe!

- It help them understand that **users may see their product with different eyes**
- It creates a **common reference point** and shared experience across the team
- Lots of observers **see more** than just one (and help with note taking)
- It's great fun! (or rather depressing sometimes!)

**Warning:** Avoid observers in the same room, it's intimidating!

# Basic tips for interviewing users

- **Demographic questions** – what kind of person are you?
- **Task questions** – how would you achieve this? What are you doing? Why? Is that what you expected to happen?

**DON'T:** ask questions that could get a yes/no answer

**DON'T:** ask leading questions (“You like this, don’t you?”)

**DO:** deviate from the script

**DO:** tell them you didn’t make the prototype. (Impartial)

- **Remember: comprehension, anticipation, satisfaction, task success**



# Think aloud

- A useful technique for qualitative interviews is to ask users to “think aloud”
- Get immediate feedback on what users are **doing**, where they **looking**, what they are **reading**, what they are **understanding**
- Some people will find this hard to do, but keep encouraging them to keep talking.
- There are two magic phrases: “uh-huh?” and “why is that?”
- While they are thinking aloud, users will be **paying more attention** to what they are doing, reducing mistakes
- Unsuitable for measuring task time at same time

# Retrospective protocol

- A different option to think-aloud is to ask users to proceed through their tasks first, then explain what their experience was like afterwards
- This means you can still time tasks accurately
- But users will still often **post-rationalise** what they did and explain away problems that they had

# Measuring performance

In usability studies you can measure many things, though the 3 fundamental things to measure are:

## **System efficiency**

...for example time to complete task, loading time

## **System effectiveness**

...for example task completion, number of errors

## **User satisfaction**

...for example preferences, uptake etc

Other aspects you can think of?

# When to ignore what users say...


The most trustworthy of user feedback is their behavior during tasks

**Beware when asking for their opinion**

- Humans **aren't good at introspection**
- It's an awkward situation: **they will say whatever they think you expect** them to
- Often they'll **blame it on themselves**
- **They aren't designers:** they don't realise the impact of their suggestions

**This means**

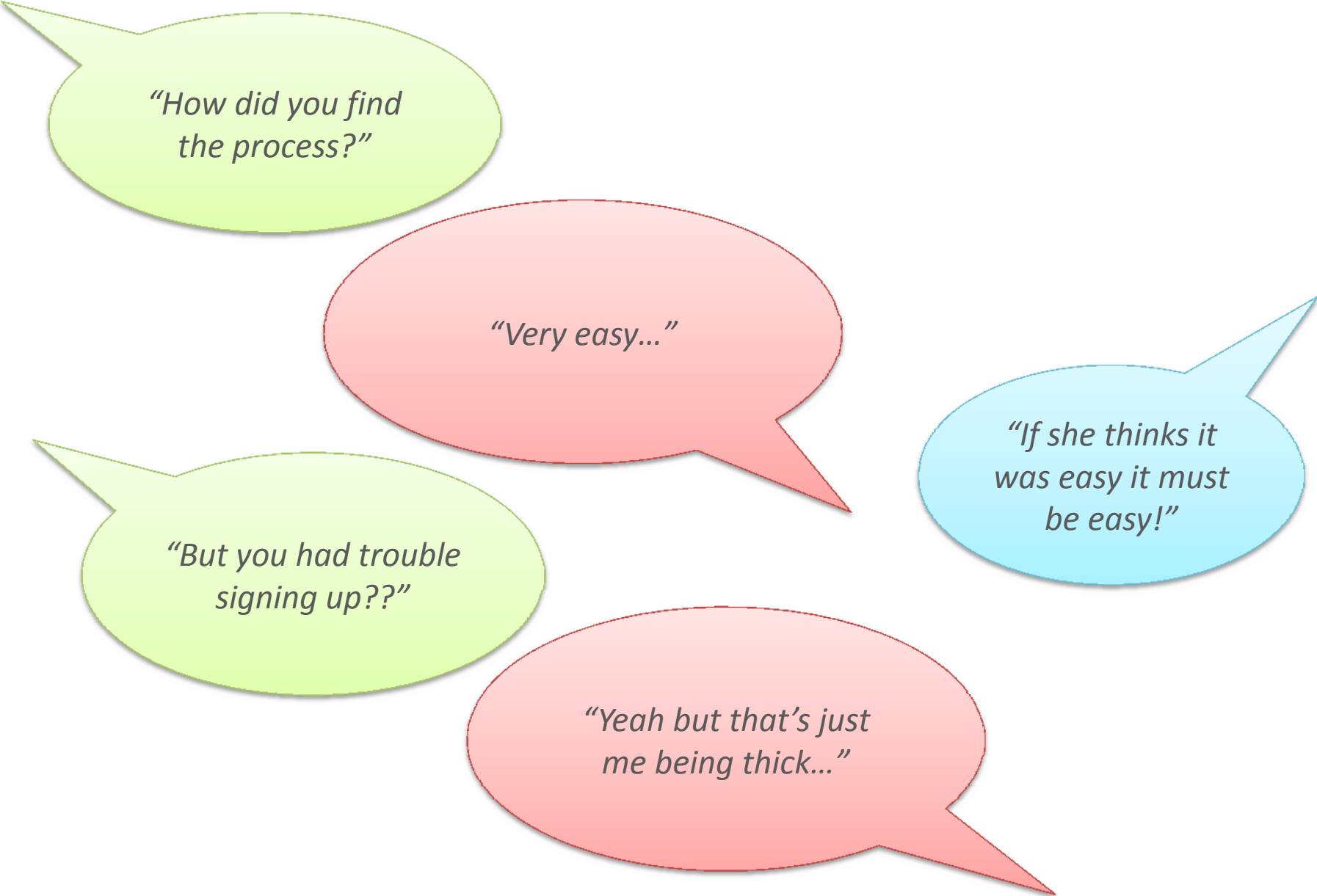
- **Avoid what-if** questions. They are not effective.
- **Always interpret** what users say carefully.



*“How do you feel  
about these security  
questions?”*

*“Uuh...there are a lot of  
questions here but I’m  
sure they are there for  
a good reason...”*

*“See the user thinks the  
procedure is  
reasonable!”*



*"How did you find  
the process?"*

*"Very easy..."*

*"If she thinks it  
was easy it must  
be easy!"*

*"But you had trouble  
signing up??"*

*"Yeah but that's just  
me being thick..."*

# Exercise

- Choose one of the phones in the session
- Set some evaluation goals
- Set your interview script to last for about 5 minutes
  - Introduction
  - Explore current user attitudes & behaviours
  - Complete some tasks (from user or yourself)
  - Wrap-up questions
- Get into pairs and participate in each other's usability test
- Report back!

# Analysing and documenting findings

- **If an issue happens to only one of the six users, is it important?**  
...It's your call
- **What was the impact?**  
Catastrophic or just an annoyance?
- **How persistent was the problem?**  
Was it a one off or would it occur all the time?
- **Problem severity is a judgment call based on your expertise in user behaviour & usability theory**
- A **spreadsheet** or a **PowerPoint** showing findings and severity rating is the most common deliverable



## *2.3.8 Eyetracking & A-B testing*

# Eyetracking

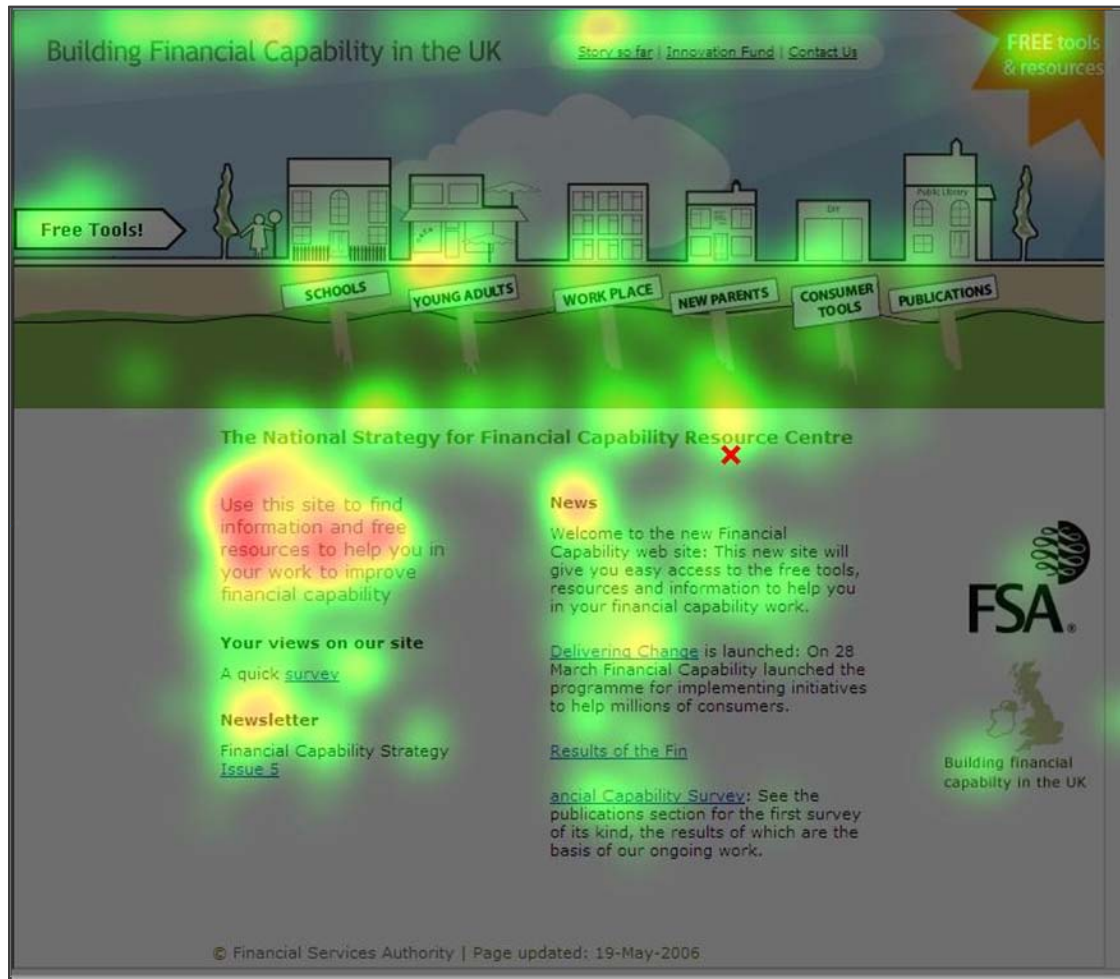
Eye-tracking records a reflection of invisible infra-red light off the cornea to **show where people are looking** on a screen.

## The software records:

- The spots where people look (**fixations**)
- The eye movements (**scan path**)
- The **length of time** people look at a specific area of interest (fixation duration)
- The **mouse clicks**
- The **pages that they are looking at**

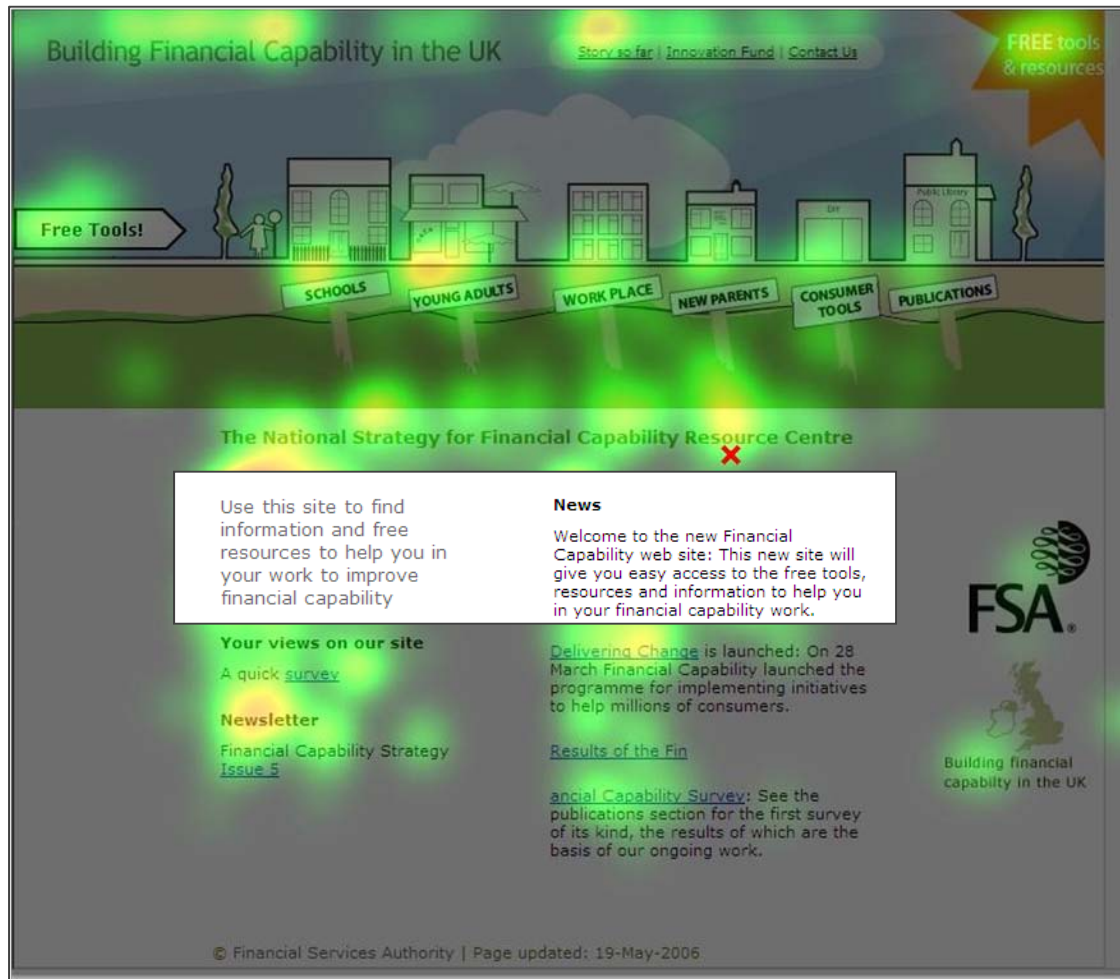
*Example: An eye tracking recording*

# Example: An eye tracking recording



***Discuss: What does this tell you?***

# Example: An eye tracking recording



# Eyetracking

## Downsides:

- Eye-tracking requires **expensive equipment, trained operators** and involves **additional set up** (at least ½ a day extra) and analysis time (at least one day extra)
- It only tells you **where users look**, but **not what they see**, why they are looking there and how they feel

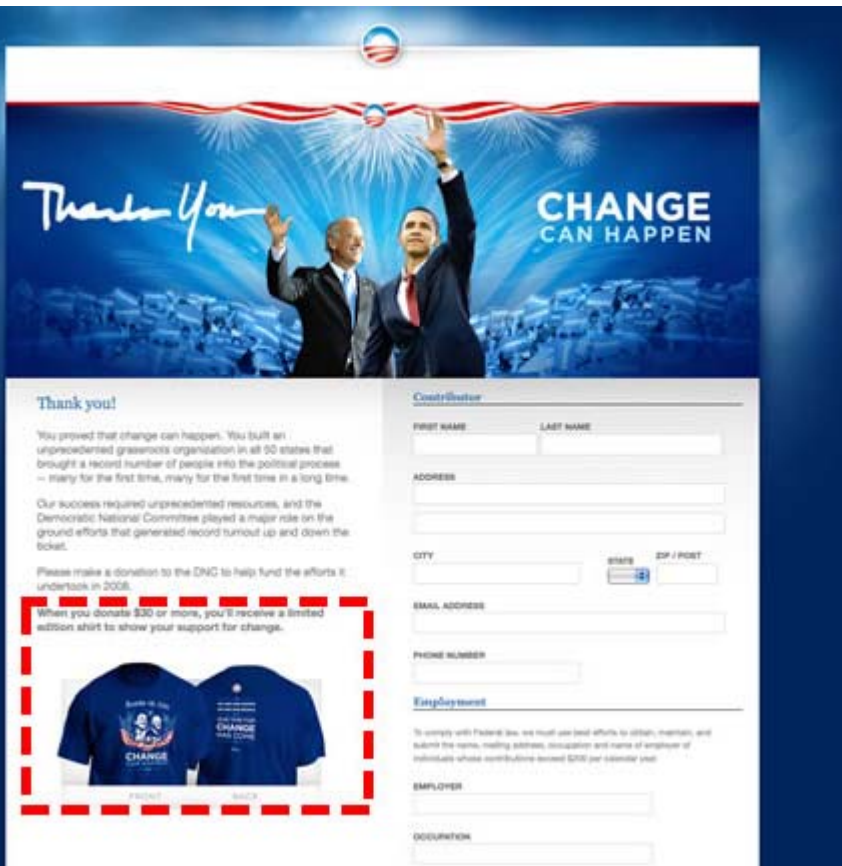
Eyetracking needs to be **done in combination with in-depth interviews** and proper user testing.

# What is A-B testing?

- A-B testing is where two versions of a website are created that differ by one controlled variable
- These sites are then shown to alternative users, and the results are analysed to see which version has better performance
- This is supported through online analytic tools like [Google Website Optimiser](#)



# Case study: BarackObama.com



When you donate \$30 or more, you'll receive a limited edition shirt to show your support for change.



Control: dark blue t-shirt

When you donate \$30 or more, you'll receive a limited edition shirt to show your support for change.



Variation A: single t-shirt

When you donate \$30 or more, you'll receive a limited edition shirt to show your support for change.



Variation B: modified lighter blue

When you donate \$30 or more, you'll receive a limited edition shirt to show your support for change.



Variation C: dual white t-shirts

Please make a donation to the DNC to help fund the efforts it undertook in 2008.

Variation D: no incentive

From [Harry Brignull's blog](#)

# ...beyond launch

The beauty of the web is that it's easy to adjust things after launch  
Use web metrics to **measure uptake** post launch

## Things you can find out:

- What users click on
  - Their way through the site
  - Where they come from
  - Drop-out rates
  - Entry and exit points
- 
- More information here: [14 free tools that reveal why people abandon your website](#)

# Summary

## Today we talked about:

- The psychological basis of usability problems and behaviour on the web:
  - Sensation, perception and gestalt
  - Memory and attention
  - Implications for design
  - User behaviour on the web
- **Usability evaluation methods**
  - How to set goals for your evaluation
  - Analytical methods: heuristics and cognitive walkthroughs
  - User methods: qualitative and quantitative usability tests
  - Eye-tracking and A-B testing

# Your project...