Usability Testing

Planning and Reporting

Notes from

Learning objectives

- Be able to develop usability testing plans
- Be able to write usability test reports
- Understand the nature of human research ethics requirements when conducting studies on humans

Plan EVERYTHING

- Use a template such as at
  http://www.usability.gov/methods/test_refine/learnusa/testplan.html

- Fill in ALL the bits – even those that are completely obvious

- Take the position that you are planning the test for 5 other people to do, each is in a different part of the world. To be useful all the participants must have the same experience.

Details

- **Product under test**
  - **Exactly what and how is it going to be tested include (as relevant) version numbers etc.**
- **Test Objectives**
  - **What's the goal? What are you planning to measure**
- **Participants**
  - **How many and what times of people are to be recruited?**
- **Equipment**
  - **Be specific, eg**
    - **for desktop specify OS, screen, processor, UI devices.**
    - **For web specify device, browse.**

Details continued

- **Test Tasks**
  - **Must be detailed and fully described (and tested). See next slide**
- **Test Procedure**
  - **Full details see later**
- **Data to be collected**
  - **Be specific**
- **Data Analysis Plan**
  - **Be specific - test it out.**

Participants

- Generally it is better if people can be themselves.
  - Occasionally you need people to role-play
- Try to match age, gender and things like first language to the target audience
- How many?
  - 10-12 will generally give you good results for a single product
  - 30+ if comparing products **and** you want statistically valid results.

Task Selection

- Utterly central to what you will learn in the usability test

- There just isn't time / resources to do usability testing on everything

- Select the tasks that are 'make-or-break' for the application

- You're looking for the *risk*
  - What's novel? What will differentiate this product?
  - If you're in a 'safe' zone where you're emulating well-established interaction patterns, then you'll learn less
    - Then again, still can be important to check that you got it right!

Task Design

- Be specific*
  - "enrol in COMPSCI345 at the University of Auckland"
  - "what lecture room(s) are COMPSCI345 lectures in?"

- Record Completion Paths
  - Step through the task yourself
    - Record different routes to successfully complete
    - Time yourself
  - Note things you think are difficult or confusing
    - You are, in effect doing a Heuristic Evaluation (*assignment 1 is pretty much this with thorough documentation*)
    - **Remember** these are notoriously inaccurate (users will surprise you)
  - **Do not** show these to participants

Task Order

- For you assignment you need to think carefully about task order.
  - Do you want to fix the order – eg everyone does cash, then credit card then txt
  - Do you want to 'latin square' – eg every variation of order (there are six)

  - What do you think the difference would be?

    - This is called the learning effect.

Questionnaire

- The easiest way to gather satisfaction data is a questionnaire

- There are several 'standard' questionnaires

  - http://www.usabilitynet.org/trump/documents/Suschapt.doc

  - http://www.w3.org/WAI/EO/Drafts/UCD/questions.html#posttest

1. What are your overall impressions of the Web site?
2. If you had to give the site a grade, from A to F, where A was exemplary and F was failing, what grade would you give it, and why?
3. Name three words or characteristics that describe this Web site.
4. What are the three things you like best about the Web site?
5. What are the three things you like least about the Web site?
6. If you could make one significant change to this Web site, what change would you make?
7. Would you return to this Web site on your own in the future? Why/why not?
8. What would entice you to return?
9. Are there materials you would like to see added to the Web site? Which ones?
10. Would you recommend this Web site to a colleague? To a friend?
11. Do you have any other questions or comments about the Web site or your experiences with it?

Questionnaire – open and closed

- Open questions (as per previous slide) give you rich qualitative data
  - Best for finding the seeds of resolutions to problems
- Closed questions allow you to quantify
  - Would you recommend this website to a friend? [Circle one] YES   NO
- Yes/No is OK, but better to use Likert scale
  - This website is easy to use: Strongly Agree Agree   Disagree Strongly Disagree
  - Converts to scores (1-4, 1-7, etc.), can report mean and other statistics and graphs
- There's a whole world to writing questionnaires; starter: http://www.terry.uga.edu/~rgrover/chapter_5.pdf

Procedures

- Don't under-estimate the practical problems
  - If you get something wrong you can lose a lot of time (and, in most situations, money) having to reschedule
    - Running a pilot or walkthrough can help here
- Have you figured out…
  - How to pay / reward participants?
  - Any catering for longer session (including your staff if the experiment is in the 'field')?
  - Do you have enough power points and battery life?
  - Enough storage capacity (e.g. on your digital camera)?
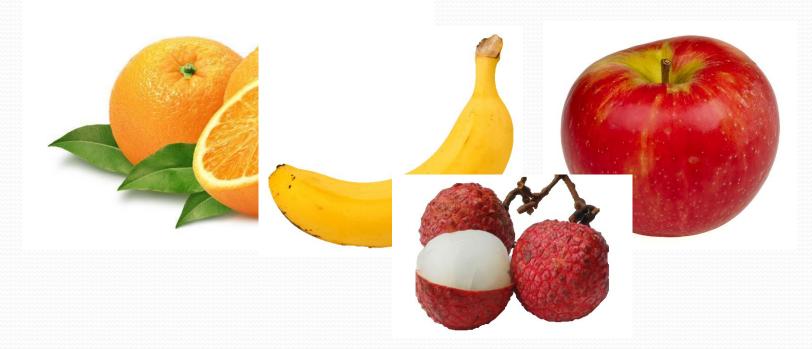 - The time and ability to 'reset' in between participants?

Details (*not* on that sample template!)

- Analysis plan
  - How are you going to turn the raw observations into assessment against your usability requirements, and into recommendations?
  - It's one thing to declare 'time' as a metric and plan to collect video, but have you defined exactly which elements of the task you are timing, and the protocol for marking the task time?
    - Is your method practical and accurate? Will it support the overall purpose of your usability test
  - What is your plan for how to report the findings
    - What, to whom, when and toward what follow-up action?

Half time entertainment

- https://www.youtube.com/watch?v=3Qg8oqTfzgU

Write a Script

- Script the usability study EXACTLY
  - Greeting
  - Ethics
  - Task instructions
  - Questionnaire

- If you don't have a script you WILL get lazy and miss instructions for later participants.

Data Collection

- **Metrics**
  - **What will you measure/collect**
    - **Video**
    - **Errors**
    - **Time**
    - **Observations**
    - **Questionnaire**
- **How many people do you need during tests?**
  - **Participant guide**
  - **Observer**

Analyse Results

- Task time and success
- Errors – you'll want counts *and* to form categories
  - Wrong navigation
  - Problems finding particular features
  - ....
- Questionnaire analysis
  - If around 10 people or less, show raw data, mean and standard deviation
  - If more than 10 people, box plots or frequency distribution graphs might be appropriate

Pilot Test

- Try the whole thing out on one or two people (or more if it's a really important and large usability study)

- After first person fix obvious problems
  - If very few corrections needed in test plan then you can go straight to testing
  - But it is much better to do a second pilot than discover major problems half way through

Analyse

- Summarize information into tables
- Use numbers where you can
- Classify comments into groups

- Run statistics as appropriate

Think!

- The big picture
- What have you found?
- What is worth fixing?
  - Is there a business case?
- How could the problems be alleviated?

Report

- Document
  - Detailed report of everything you have found
    - Three formats here
      - http://www.usability.gov/templates/index.html
    - Remember numbers are very convincing, compare:
      - Several people had trouble finding the shopping basket
      - 3 out of 7 people abandoned the task because they couldn't find the shopping basket. For the other 4 the average time to find the shopping basket was 3.59seconds (longest 8.0 seconds)
- Video
  - Imagine clipping together the 7 people looking for the shopping basket icon … with puzzled looks on their faces!

Ethics

- If you are doing a study with living (human or animal) participants in a university you will probably need ethics approval
  - Can be quite a lot of paperwork, and takes a while to get an answer (which is usually to revise and re-submit!)
  - You will need such approval for a study to be part of your dissertation or thesis
  - Many journals require such approval to publish
- Quite a few companies have similar requirements
- This is why for your assignment you are not testing on others – though you could ask classmates…

Research ethics basics

- Informed consent
  - Participant knows what they are 'in for'
    - Task, time, why you're doing it (even though you may be allowed to 'deceive' them about some aspect of the task)
    - Confidentiality of their data
    - Compensation (if any)
- Participant is clear that they are not compelled to participate
  - This is a bit of a trick in lecturers experimenting on their students! (or doctors on patients, or bosses on their employees)
  - They need to know that they can refuse, or withdraw (even retrospectively!) without jeopardising the key service (healthcare, education, employment)
- Anonymous questionnaires, esp. in public, are probably the easiest from an ethics perspective

Ethics application

- Explains protocol and goals: essentially like a test plan
    - And so it's helpful to complete one because it acts as a check on your plan
- Particular focus on issues such as who has access to the data and the risk (and benefits, if any) to participants
- Research organisations (University, District Health Board) have standing committees to review applications
    - Have representatives from a range of perspectives: clinical, legal, statistical (and Maori in NZ)

Professionalism

- Treat participants with respect
  - Assume they are not idiots, it is the software that is wrong
- Treat developers with respect
  - They may have put their heart and soul into the product and worked overtime to get if finished for you to pull it apart
- Make sure your report is
  - Fair and accurate
  - Tidy
  - Free from grammar and spelling errors

In the real world

- If you can't do a 'real' usability test
  - Get your mates, Mum, Dad, Aunty Flo to try it

- Tune-in to your own usability experiences
  - Note what was really easy
    - that's a sign of good usability
  - Note what is annoying you
  - Note when you are trying to do something you have done before and can't remember how.

Summary

- Evaluate usability early and often in development and [preferably staged] roll-out
  - Also evaluate alternatives before making a decision to purchase/adopt a system
- You need a complete and detailed testing plan
- Heuristic evaluation is a handy intermediate level between just asking a couple people for feedback and doing a full-blown usability study