

Guerilla Usability Testing

by Andy Budd

Before we start, if you have an Intel Mac laptop could you please stand up.

What is usability testing and why is it important?

Usability testing generally involves setting a series of **tasks** for people to complete and **noting any problems** they encounter – *It's as simple as that!*



Usability testing is incredibly simple. At it's most basic we set a series of tasks for our users to complete and then note any problems that occur.

We do this because users are unpredictable



We do this because users are unpredictable

We like to think we know who they are...



But we end up designing for people like us



But invariably end up designing for people like ourselves.

We just can't help it. It's not because we're bad designers. It's just human nature.



We design sites assuming that users will interact with our site in a rational, logical order.

That they know what they want, will take time to understand the interface and weigh up the options, before selecting the optimal solution.



But this is crazy. It's not how we ourselves use the web, so why do we assume other people do?

In reality users attention will dart around the page looking for the first thing that jumps out at them. By bombarding them with pointless offers, they have become adapt at ignoring anything that looks like an advert or marketing copy.

We think people use our sites like a game of chess. However most of the time it's more like a game of snap.

I've moderated countless sessions where users have said something like 'I would have expected their to be X' where X was right under their mouse pointer. They just hadn't noticed.

It would be easy at that stage to blame the 'stupid user' but in fact it's the fault of the 'stupid designer' for not having a better understanding of basic user behaviour

People are **very quick** to associate their online experiences with your offline brand. It's called transference.

If your site is a pleasure to use, they will associate those **positive feelings** towards your brand.

If your site is difficult or frustrating to use, they will associate **negative feelings** with your brand. At best they will think you're sloppy and don't care. At worst they will tell all their friends how terrible you are.

Usability is important for a lot of reasons, but you guys know this so I'm not going to bother lecturing you.

Suffice to say...



A classic example for me is the Iberia Airlines check-in process.

See if you can find out where to check-in on the Iberia Airlines check-in page.

This is meant to make the process of flying easier, but it actually makes it more frustrating.

Undoubtedly leads to more airport check-ins and additional cost for the airline.



Did you guess correctly?

It took 3 usability consultants a good 5 minutes to figure this one out.

Benefits of usability testing for designers

- Help inform the design process
- Uncover unknown or unexpected issues
- Explore users mental models of the world
- Fix problems early in the process when they are easy to fix
- Test assumptions
- Highlight unnecessary features
- Provide objectivity help and solve opinion battles
- Set baselines and measure improvement
- Get stakeholder buy-in



Benefits of usability testing for site owners

- Fix problems early in the process when they are cheaper to fix
- Improve customer satisfaction and retention
- Increase conversion rates
- Reduce maintenance, training and support costs
- Get stakeholder buy-in
- Reduce project risk
- Set baselines and measure improvement
- It's so cheap you'd be stupid not to



So how do I get my boss to agree?

* If you present something as an issue, management will often have a 'wait and see' attitude. So just do it anyway.

It's always easier to ask for forgiveness than permission

* Set up a quick and dirty test and show how useful it is

* Show how usability testing will reduce the chance of bad decisions causing the project to fail. And if the project does fail, usability testing can be your 'get out of jail free' card.

* Estimate how much the testing will cost versus how much it could cost the company in lost sales, support costs, maintenance, training. Cost-benefit will work on some, but not all managers.

* Show the benefits for your team. i.e. This would really help bob the developer!

- Don't ask permission. Just do it anyway.
- Show, don't explain
- Explain how it will reduce the risk that the project will fail.
- Demonstrate the return on investment.
- Highlight the benefits for your team.
- Explain how it will provide hard data you can measure success against.

So how do I get my boss to agree?

* If you present something as an issue, management will often have a 'wait and see' attitude. So just do it anyway.

It's always easier to ask for forgiveness than permission

* Set up a quick and dirty test and show how useful it is

* Show how usability testing will reduce the chance of bad decisions causing the project to fail. And if the project does fail, usability testing can be your 'get out of jail free' card.

* Estimate how much the testing will cost versus how much it could cost the company in lost sales, support costs, maintenance, training. Cost-benefit will work on some, but not all managers.

* Show the benefits for your team. i.e. This would really help bob the developer!



Many people equate usability testing with QA.

It's something you do at the end of the project to make sure you've done everything right. It's often an ass covering exercise just to check the boxes, and there is little desire (or ability) to make changes at this late stage.

Summative testing is essentially evaluating what you've already done to make sure it's been

done correctly.

However I'm more interested in formative usability testing. This approach sees usability as part of the design process. You use testing to provide insight and the resulting knowledge helps you 'form' the final product.

Iterative design vs. genius design

With formative testing the results are used by the designers and fed back into the design process. With summative testing a report is written and usually left to gather dust. Useful for baseline comparisons and metrics, but little else.

The number of times I've spoken to clients who did usability testing a year ago, but haven't actually implemented any of the results. I guess at least they are aware of the problem.

Formal Testing

If you've been involved with a usability test before I imagine it was probably a more formal usability test.

These tests tend to be carried out by large, specialist companies with dedicated test suites. These suites often have one way mirrors, live audio and video feeds, and notice boards for collaborative note taking and discussion. Quite often they will also have some form of eye tracking (which I'll come onto in a second.)

They will use a professional facilitator/HCI person and generally test on a larger sample for more statistical rigour. Tests tend to be more quantitative than qualitative and the output is usually in the form of a written report to the site owners or developers.

This set up is great for getting the whole team together to watch the tests.

- Will recruit test subjects for you
- Facilitators less likely to bias results
- Comfortable facilities
- The whole team can be present
- Quantitative data *likely* to have more statistical validity
- Professionally written report or presentation
- Time consuming to set up, run and analyse results
- Expensive so often only run once at the end of the project
- Facilitators only have shallow knowledge of the project
- Focus on quantitative over qualitative reduces insight
- Smaller test samples can invalidate quantitative results
- Most projects don't require the level of statistical rigour
- Easy for the results to be 'thrown over the fence' and forgotton

- Will recruit test subjects for you
- Facilitators less likely to bias results
- Comfortable facilities
- The whole team can be present
- Quantitative data likely to have more statistical validity
- Professionally written report or presentation



- Time consuming to set up, run and analyse results
- Expensive so often only run once at the end of the project
- Facilitators only have shallow knowledge of the project
- Focus on quantitative over qualitative reduces insight
- Smaller test samples can invalidate quantitative results
- Most projects don't require the level of scientific rigour
- Easy for the results to be 'thrown over the fence' and forgotten

Cons



Eye-tracking uses specialist hardware and software to track the movement of the eye on the page. Uses a mixture of video cameras and infrared light to compute gaze direction.

This information can be very interesting, especially if you're testing things like the visibility of buttons etc. However eye-tracking can be a costly process and it's very difficult to interpret, so is mostly used in academic studies at the moment.

This is a Tobii eye tracker. It costs around £30k. It beams infrared light into the eyes to see where they are fixating on.

eye trackers are admittedly cool bits of kit and make you feel like you're doing scientific research.

Managers especially love them as seemingly you can't argue againts science



They produce beautiful heatmaps like this which look sciency so they must be good.

The problem is, eytracking in general and heatmaps in particular are open a good bit of interpretation. So it's very easy to use them to "prove" something that may actually not be the case. As such, many agencies will pick the heatmap that best matches their assertions.

They are good for pulling out general trends so I think there is some validity for general, academic research purposes when you're looking at behaviour across a number of sites.

I'm yet to see heatmaps used to find a usability problem that could not have been found by other means



Gaze maps are potentially more useful, but again they are difficult to analyse well and take a lot of time

However they are useful for things like confirming is users actually noticed something like a button, feedback message or advert.

However with so much low hanging fruit available on the web, I think eye tracking is overkill for 99% of cases. It's like driving your car by looking at the dashboard. You're so focused (even overwhelmed) by the stats, it's easy to miss where you're going!

The cost to benefit ration just doesn't seem to be there to me.

I find formal testing is best for summative testing at the end of a project



Gaze maps are potentially more useful, but again they are difficult to analyse well and take a lot of time

However they are useful for things like confirming is users actually noticed something like a button, feedback message or advert.

However with so much low hanging fruit available on the web, I think eye tracking is overkill for 99% of cases. It's like driving your car by looking at the dashboard. You're so focused (even overwhelmed) by the stats, it's easy to miss where you're going!

The cost to benefit ration just doesn't seem to be there to me.

I find formal testing is best for summative testing at the end of a project

Informal (Guerrilla) Testing

Screenshot of a guerilla test http://www.flickr.com/photos/raphaelquinet/513349541/

You don't actually need much equipment to run a usability test. Just a computer to run the test, somebody to moderate and somebody to take notes.

Instead of a note taker, I prefer to set up a video recorder pointed at the screen. This captures the mouse movements as well as the audio, so you don't need to take notes.

Alternatively you could have two videos, one recording the screen and one recording the person. Or a video and some screen capture software. However you then need to composite both feeds and this can get complicated.

Guerilla testing tends to be carried out anywhere possible, from an empty meeting room to a local cafe. It's usually (although not necessarily) carried out on a one-to-one basis by a member of the design team.

Instead of focusing on quantitative data, these tests usually focus on qualitative data for design insight rather than statistical validity. The results are usually fed back into the design process. As such, they tend to be formative rather than summative.

- Very quick and easy to perform
- Relatively inexpensive so can afford multiple tests
- Qualitative nature can provide improved design insight
- Results can be fed back into the design process almost immediately with no loss of signal
- Have to recruit subjects, arrange location and write tests yourself.
- Familiarity with project may introduce personal bias
- Less statistical rigour on quantitative results

- Very quick and easy to perform
- Relatively inexpensive so can afford multiple tests
- Qualitative nature can provide improved design insight
- Results can be fed back into the design process almost immediately with no loss of signal



- Have to recruit subjects, arrange location and write tests yourself.
- Familiarity with project may introduce personal bias
- Less statistical rigour on quantitative results





Remote, moderated testing

Don't have to have people come into your office. You can go into theirs (virtually) and experience the frustrations they face.

Quite a few people are doing remote usability testing.

There are expensive dedicated services like UserView, but you can also use net meeting or screen sharing software like Co-Pilot.

Not done it myself, but looks promising.



Home > Research Methods > Online Usability Testing

RESEARCH METHODS

Overview

Online Usability Testing

Online Surveys

UX Benchmarking

Prototype Testing

Online Card Sorting Studies

Voice of Customer Solution

Search Experience Studies

Mobile Interface Testing

International Usability Studies

Online Usability Testing

Haven't gone remote yet?

We offer UX pros the ability to run online, taskbased studies with geographically dispersed participants. Users participate simultaneously in their natural context and no human moderation is needed.

This method is called 'unmoderated remote usability testing'. With it you get objective statistics to *quantify* usability, honest feedback, usability metrics + web analytics combined.



How Does It Work?

Conduct task-based studies over any web-based interface with hundreds of users participating simultaneously from their home or office. Get quantifiable measurements of usability and user experience.

You get direct, honest feedback from statistically significant samples. It's very cost-effective since we test users online, allowing for both local and

international testing, and it's a great complement to other research methods.

What Do You Get?

UZ SSE gathers rich user experience data and allows researchers to perform advanced analysis. Our cutting-edge technology is capable of gathering both qualitative and quantitative data, such as:

Effectiveness ratios

Efficiency ratios (time and clicks)

- Visual clickstream paths
- Click mapping (clicks per page)

Remote, unmoderated testing

Satisfaction & perception indicators



Want to know our pricing models?

Also interesting for you

- ► Case Studies 👗
- Check out our webinars!
- Lab + online usability testing

Keynote, UserZoom



Not focus groups

http://www.flickr.com/photos/twschulz/1598096681/

Well meaning managers can get it into their heads that usability testing is the same as a focus group because, well, it involves talking to users doesn't it.

Don't let your manager or client run a usability test because they have done a few focus groups, it sounds like fun and they want to cut corners. Not without a serious amount of training.

Observation based research is very different to opinion based research and a lot more valid. How many of you washed your hands after going to the toilet?

Focus groups are great for learning about peoples attitudes, beliefs and opinions or their reactions and ideas about a prototype. However they can't tell you how people actually use the prototype or what problems they will encounter.



Planning your test

As with formal testing, you should start with a test plan. However this can be a simple 2-3 page document

that outlines What, Why, When, Where, How

- Introduce the project
- State the goals of the test
- Profile your desired test subjects
- Create a test subject screener
- Outline your testing methodology
- Define your tasks
- Write a testing script
- Design a post test survey

Planning your test

Select your tasks

You can't test everything so determine which are the most important tasks to test. Test the most common tasks, the most complicated tasks and tasks you think may have usability issues. Create a believable flow of tasks. e.g. register > complete profile > upload photo etc. Start with simpler tasks and slowly increase complexity Base tasks on your user scenarios. Select around 5-6 tasks that should take approximately 45 min.

Run through the tasks to ensure they are achievable. No point setting tasks you know will fail!

Good and Bad Tasks

Bad: Search for a bookcase

Good: You have 200+ books in your fiction collection, currently in boxes strewn around your living room. Find a way to organize them.

Don't load tasks with the solution to the design. It's much harder than you think.

Instead create realistic scenarios

а and spend a day ing with clay!

clear left

Earn £20 for an hour of your time

We're looking for a few people to help us test a new website we've been working on. You don't need to be technical or understand anything about web design. You just need to know the local area and enjoy the

We'll send you a quick survey to fill out, and if you meet our criteria, we'll arrange a time for you to pop in. And as a way of thanks we'll give you £20 for an

We're looking for a few people, so if you or any of your friends are interested, please drop Sophie an

s individually or image

ue roller blind. derful holiday photo, nting, a beautiful s are endless.

nd samples: 57 819 5 ODA ativerollerblinds.co.uk

Recruiting Test Subjects

- Client contacts
- Via the client website (ehtnio.com?)
- Recruitment agency
- Market research agency
- Twitter
- Local notice boards and mailing lists
- Other offices/departments
- Classified ads e.g. gumtree

All your subjects should be from the same user type of persona These should be based on behaviour rather than demographics e.g. For a travel site you may have business traveller, tourist and travel agent.

Pay them. £20-50. Vouchers, tickets, wine and chocolates etc. Feel free to be creative!

In a survey of usability companies only 36% used external recruiters, and even then this was often to supplement their own recruitment. only 9% always used external recruiters. Takes about an hour per person recruited.

Client contacts

- Via the client website (ehtnio.com?)
- Recruitment agency
- Market research agency
- Twitter/Facebook
- Local notice boards and mailing lists
- Other offices/departments
- Classified ads e.g. gumtree

s individually or image

ue roller blind. derful holiday photo, nting, a beautiful s are endless.

nd samples: 57 819 5 ODA ativerollerblinds.co.uk

Recruiting Test Subjects

All your subjects should be from the same user type of persona These should be based on behaviour rather than demographics e.g. For a travel site you may have business traveller, tourist and travel agent.

Pay them. £20-50. Vouchers, tickets, wine and chocolates etc. Feel free to be creative!

In a survey of usability companies only 36% used external recruiters, and even then this was often to supplement their own recruitment. only 9% always used external recruiters. Takes about an hour per person recruited.

How many subjects are enough?

There is no easy answer here.

Several people including Jakob Neilsen have suggested that testing on 5 users will uncover 80% of the problems on your site that have a high likelihood of detection.

This is based on binomial probability distribution.

After 5, you get diminishing returns.

However these must be five of the same type of user.

In my experience, 6 is a good number. You can run them all in a single day and you're still OK if somebody drops out.

Laura Faulkner of the University of Austin did an analysis of the results of a test using 60 subjects. 5 test subjects did find an average of 85% of the total issues found. However the percentage of problems found by any one sample set ranged from 55%–100% Groups of 10 found 95% of problems with a min of 82% 15 found 97% with a min of 90% 20 found 98% with a min of 95%

If you want to discover the majority of issues, 10-20 participants is ideal.

What do we test?
SEARCH ESSAVIS P SEARCH ESSAVIS P SEARCH AU SEARCH AU ESSA 75 UN Low fidelity paper prototypes

A low fidelity paper wireframe

For all paper prototypes, you will want to encourage people to 'click' on the paper. Probably a good idea to have somebody play the role of the computer, providing open dropdowns, etc.

A little bit artificial, but in my experience people are good at suspending disbelief.

Can laminate paper prototypes to subjects can write on them.

5. Exhibitor Zone



A medium fidelity paper wireframe



Screen name did something

5 minutes ago Screen name did something

6 minutes ago Screen name did something

1 hour ago Screen name did something

3 hours ago

Add a new goa

A medium fidelity HTML prototype

Your groups

Group name

Group name

Your goals

HTML/CSS prototypes

Could easily do a whole workshop of wireframes and prototypes. In fact, we're thinking about doing this very thing.

For more information, check out the session Rich and James did at SXSW

• <u>States low-energy lightbulbs throughout my home</u>

Use public transport for my journey to work

Universal Channel Diva Universal Studio Universal Syfy Universal 13TH STREET

UNIVERSAL NETWORKS



You can produce fairly realistic feeling sites by slicing and dicing comps in Fireworks/ photoshop and setting up hot areas.

Your Saved Ads (0) Manage Your Ads



Or even the final site, although it's often too late by this point!

Inverness

Take Home: It doesn't matter what you test on, as long as you test early and (preferably) often.



The basic set-up

http://www.flickr.com/photos/balakov/1485067045/

You don't actually need much equipment to run a usability test. Just a computer to run the test, somebody to moderate and somebody to take notes.

Instead of a note taker, I prefer to set up a video recorder pointed at the screen. This captures the mouse movements as well as the audio, so you don't need to take notes.

Alternatively you could have two videos, one recording the screen and one recording the person. Or a video and some screen capture software. However you then need to composite both feeds and this can get complicated.



Expensive (starts at \$1,500) and difficult to set up. PC only, although possible to get it working in parallels More suited to traditional lab work than guerilla usability testing

Our goal was to simplify the voter registration process during the Obama campaign. Silverback was clearly the easiest way to gain true insights into user behavior, and resulted in a successful, more informed site at launch. There is nothing silver about Silverback, all gold in my opinion!

so says Scott Thomas of Obama for America | Next testimonial »

Clearleft presents

Silverback

Spontaneous, unobtrusive usability testing software for designers and developers.



Silverback

Cheap. \$69.95.

Very easy to set up and run.

Basically records the screen, your mouse movements, audio and video feed from isight at the same time.

Can be controlled by the apple remote. Use it to set markers in the video.

Selecting a moderator

Some people worry that if you choose somebody from the team they may not be objective. And it's true, if you pick somebody who is invested in the current outcome and doesn't want it to change, you'll get biased results.

However usability testing is all about insight. So I think it's important to use a moderator who

has a thorough knowledge of the project and can use this to gain insight into the users state of mind and how the design can be improved.

If you have a good UX person, they should love to get contact with the users to learn what they are ding wrong and how they can do it better.

Exercise 1: pick a site

- Lastminute.co.uk
- Easyjet.com
- Firstchoice.co.uk
- Eurostar.com
- Kelkoo.co.uk
- Pricerunner.com
- Currys.co.uk
- Universal-playback.com

- Compairethemarket.com
- GoCompare.com
- Moneysavingexpert.com
- Confused.com
- uSwitch.com
- Ticketmaster.com
- Moonpig.com
- interflora.co.uk

10 minutes

In your groups pick a site that you're familiar with but preferably not involved in.

Choose a site that you suspect will have a number of usability problems.

You're looking for a site that has a couple of fairly well defined user paths. So buying a ticket, booking a hotel, comparing the price of a camera or buying a washing machine.

Exercise 2: choose your goals

Identify the key user journeys on the site and set some goals for things you want to test.

Broad: we want to understand how people use our service to compare multiple items.

Intermediate: we want to test how easy it is for people to use our booking engine so we can fix any usability problems we find. **Specific**: we want to understand why so many people drop out of our check-out process, so we can increase conversions by 30%.

10 minutes

Exercise 3: set three tasks

- The tasks should represent common user goals.
- The tasks should support your testing goals.
- The wording should not hint at the solution.
- All 3 tasks should be closed-ended.
- You should be able to test all 3 tasks in 15 min.

15 minutes

All three tasks could easily be supporting a single testing goal.

Exercise 4: review neighbours tasks

Pass your task sheet to the next table.

Discuss the tasks you were given as a group and review their suitability. Update if necessary and pass back.

During the break, review the changes and do a quick dry run. You'll be doing it properly when you get back.

15 minutes

Running your test

FLO

ORA : Vitte!

A

- Don't keep people waiting around
- Have everything already set up and waiting
- Welcome people and thank them for their time
- Explain what you'll be doing and how the test will work
- Remunerate them in advance so they understand it's not contingent on their performance
- Ask them to sign a waiver
- Remind them that you're testing the system, not them
- Encourage them to think aloud

Running your test



Think aloud

Think aloud testing (introspection) is usually the type of testing we perform as it's the easiest to get insight out of. However be aware that by asking users to 'think aloud' you're changing their normal pattern of behaviour. For instance 'time on task' calculations may no longer be valid.

That being said, you'd be amazed by how tolerant people are of this type of test and their willingness to suspend disbelief and play along.

Be aware that thinking aloud may make people more conscious and therefor attentive to what they are doing. Ironically it may also distract them from what they are actually doing. Subjects often go quiet when they get stuck.

The other option is retrospective introspection when you let people do a task and then get them to explain what happened after the fact.

What a usability test looks like





Think aloud exercise

A good way to explain to people how 'thinking aloud' actually works, is to get them to do a practice run on a familiar object.

Sometime what I do is get people to set the alarm on their phones or change the timezone, and have them "think aloud"



Be a good moderator

Sit behind and slightly to one side of the subject so they can't see you. Start them off simply, maybe by just asking them to explain what they think the site is about. Set one task at a time. Don't force them to remember details

Don't care about their ideas or what their mum may think

- Be friendly and put the person at ease
- Stay objective and detached. Try not to bond
- Start with a simple task
- Keep instructions simple
- Don't give hints or ask leading questions
- Give encouraging but non-committal feedback
- If you don't understand an action feel free to ask why they did something
- Know when to stop a task but don't rescue people too soon
- Summarise key findings

Be a good moderator

Sit behind and slightly to one side of the subject so they can't see you. Start them off simply, maybe by just asking them to explain what they think the site is about. Set one task at a time. Don't force them to remember details

Don't care about their ideas or what their mum may think

- "What are you currently thinking?"
- "Describe the steps you're going through here"
- "What do you think will happen next?"
- "Is that what you expected to happen?"
- "That was confusing?"
- "And you were expecting ...?"

Moderation techniques

If people go quiet, prompt them. Don't say "what is your problem?" say "what are you thinking?"

Answer a question with a question

Echo back comments with a question mark

Let statements tail off into an unanswered questions

Be a good observer

I've moderated countless sessions where users have said something like 'I would have expected their to be X' where X was right under their mouse pointer. They just hadn't noticed.

It would be easy at that stage to blame the 'stupid user' and move on, but users aren't that stupid.

if anything it's the fault of the 'stupid designer' for not having a better understanding of basic user behaviour

We're very quick to judge the work of people in other fields, but slow to judge our own.



Usability testing mistakes

- Not doing a proper run through of the test in advance
- Incomplete or buggy prototypes
- Not scheduling enough time between sessions
- Bad task design
- Accidental Revelation
- Getting frustrated or losing detachment

Usability testing mistakes

Wrapping up

- Answer any questions they may have
- Administer the post test survey
- Thank them for their help
- Show them out
- Capture any important notes immediately
- Prepare for the next test



Exercise 5: run test three times

- Chose roles: moderator, observers, test subject.
- The test subject moves to the next table.
- Tests should last 15 minutes.
- Swap roles. Lather, rinse, repeat.





If you're doing it on your own, a data capture spreadsheet is probably best.

If you're doing it in teams (live), sticky notes are a good option

Θ	00											usa	bility-te	sting-d	ata.xls										
2			. 8		Ê.	₫.	S .	പ -	Σ.	ول	<mark>⊼</mark> ↓		10	100%	•										
New	Open	Save P	rint Impor	t Copy	Paste	Format	Undo		AutoSum	Sort A-Z	Sort Z-A	Gallery	Toolbox	Zoom	Help	,									
										Shee	ts	Charts	Sm	nartArt G	raphics	-	WordA	rt							
\diamond		A		В		C		D		É	F		G	1	H		I	J [K		L	M	N	0	
1	Yellow Submarine User Testing																								

5 Test Subjects 6 Participant 7 1 8 2 25/01/1968 10:00 Jon Paul 25/01/1968 11:00 9 3 25/01/1968 12:00 George 10 4 Ringo 14:00 25/01/1968 11 5 Old Fred 15:00 12 13 Task List 14 Task Description 1 Start Yellow Submarine 15 2 Journey to Pepperland 16 Imiate Seargent Pepper's Lonely Hearts Club 3 Band 17 Beat the Blue Meanies 4 18 Have a big party 5

2 3

4

19

20 7

6

Task Success Key							
Menu Label	Description						
Easy	1st or 2nd try - no problem						
Medium	1st or 2nd try - observed difficulty						
Hard	2nd or 3rd try - expressed difficulty						
Assist	Succeeded with assistance						
Fail	Failed or gave up						

Issue Severity Key								
Menu Label	Description							
Minor	Causes minor problems							
Moderate	Causes moderate problems							
Severe	Causes significant problems							
Critical	Causes major problems							

26			
28			
29 30	Data logger) 4 i ¥
	IA A PS Data +	(- × × (
	Normal View Ready	Sum=0	

Performance Metrics

Completion Rates – Are users able to complete the task?
Time on Task – How long does it take users to complete?
Page Views – How many pages does it take to complete?
Errors – The number and severity of errors per task?
Satisfaction – How does the user rate the system?

You'll be surprised how many people actually fail tasks in a usability test.

Can also collect other things like percentage of time on the correct path, learnability over time etc.

Great for a scientific study, but a bit of a waste of time in a guerilla test Quantitative more expensive than qualitative.

Need more subjects for statistical validity (20+) and more time analysing results Need to test against a benchmark e.g. the old site or a competitor site

Review and implement ASAP

Review the videos while the test is still fresh in your head Capture the data in a spreadsheet Patterns will start to emerge Note the problems and your recommended fixes Summarise your findings in report or presentation (shoe example)

Compile a short 'highlights' video Make the changes ASAP



Rapid iterative testing and evaluation.

A technique formalised by Microsoft while developing the training level for Age of Empire II. Rather than run a series of tests then analyse the results and make changes, the MS team would analyse the results between tests and make obvious changes before the next subject. This allowed them to

330 8 - - - - 8

- * Fix obvious issues quickly then test the fix
- * Avoid having the same issues come up time and again
- * Discover new issues that the unfixed issues would have hidden

Used 16 participants over 6 iterations. 31 problems were discovered and 30 fixes made and 6 re-fixes.

This is most appropriate when you want to find and fix issues as quickly as possible, rather than ensuring that every possible issue has been found.

Reports and Presentations

An informal discussion of the findings is best

If you're going to write a report, keep it short and sweet, otherwise it won't get read. Include an executive summary on the first page. Classify the severity of problems.

Include positive findings as well.

Exercise 6: KJ sort your findings

- write down each issue you found on a sticky note.
- once you have a handful, stick them on a free wall.
- don't stop until all the issues you saw are on the wall.
- start grouping the issues into atomic units.
- give each group a name.
- individually, pick the 5 most important issues.
- number them from 5 (most important) to 1 (least important)
- tally up the results next to each issue

Exercise 7: present your findings

Present your findings and recommendations to the group.



10 minutes
If you buy one book...

The how-to companion to the bestselling Don't Make Me Think! A Common Sense Approach to Web Usability



The Do-It-Yourself Guide to Finding and Fixing Usability Problems

Really old, but still a classic despite what you may think of Jakob

If you buy two books...

Jeffrey Rubin and Dana Chisnell Foreword by Jared Spool

Handbook of Usability Testing



How to Plan, Design, and Conduct Effective Tests

SECOND EDITION

Really old, but still a classic despite what you may think of Jakob