# Lecture 6

Evaluation and Usability Testing

FIT5152 - User Interface Design and Usability

Learning objectives

- Understand what evaluation means in HCI
- Learn about different evaluations options appropriate to the design phases
  - Expert reviews
    - Heuristics evaluation
    - Cognitive walkthrough
  - Usability testing

Requirement gathering and Design

(User and Task analysis)

Evaluation

Prototyping
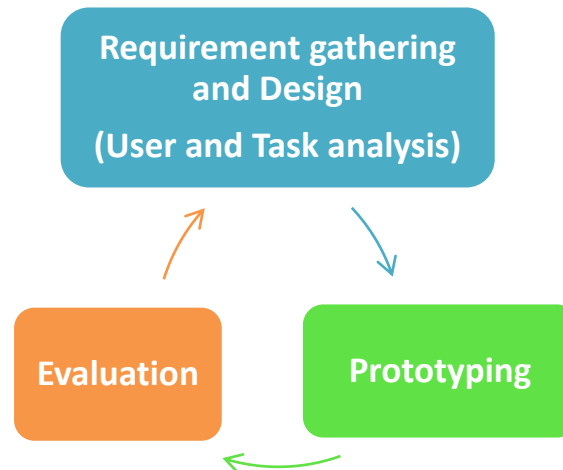
HCI process (Adapted from Greenberg, 1996)

What is Evaluation?

- Evaluation involves using different methods to test and assess a product, and solicit feedback.
- Evaluation can be conducted at different stages of design.
- Evaluation can be conducted for new or existing products.
- The purpose is to check and ensure if a product meets the user needs, if there are any issues, and whether users like the product.

Evaluation Aims

- Are the options and controls visible to the user?
- Does the user know how to interact with the UI elements?
- Is the navigation and workflow appropriate?
- Does the product do what the user expected it to do?
- Is the functionality correct from the user's perspective?
- What are the good and poor features?

Why should systems be evaluated?

- Extensive testing is a key to successful design.
- To understand how real users use the product and if it meets their needs.
- To identify the issues and addressing them.
- To reduce user errors.
- To improve user experience and increase user satisfaction.
- To ensure the product meets the standards and guidelines.
- To compare with similar products.
- It is cheaper to fix a problem earlier in development than later.

Different Types of Evaluation

- **Who** will do the evaluation?
  - Users or experts? Who is available? Who has the skills?
- When evaluators are **experts** (Expert Reviews):
  - Cognitive walkthrough
  - Heuristic evaluation
- When evaluators are **users**:
  - Usability Testing
  - Field studies
  - Thinking aloud

Different Types of Evaluation

- What type of evaluation?
    - Depends on the design phase and the type of the product
    - Depends on the criticality of the system
    - Depends on the resources and time

    For example for usability testing, you need to:
    - have access to users,
    - invite them through the right channels,
    - consider an appropriate time and location,
    - obtain ethics approval from involved organisations,
    - design data collection methods (instrument) such as interview and questionnaire questions,
    - determine how you would analyse data and report them

(Shneiderman et al. 2014)

Expert Reviews

- Expert reviews involve the evaluation of a product to identify usability problems and provide recommendations
- Who are the experts?
  - designers
  - usability experts
  - or those who have extensive knowledge about the user, the product, or the technology
- When to do expert review?
  - Can be done in early or late stages of the design
- The output of expert reviews
  - A formal report with identified problems and recommendations

(Shneiderman et al. 2014)

Issues with Expert Reviews

- Even knowledgeable experts might not know how real users, particularly first-time users, will interact with the product.

- "The danger with expert reviews is that the experts may not have an adequate understanding of the task domain or user communities."

(Shneiderman et al 2014, pg 136)

Cognitive Walkthrough

Cognitive walkthrough is usually used iteratively in the early stage of design

The preparation phase:

- Experts need a task description
  - Tasks need to be carefully selected (e.g. high-frequency tasks)
- Experts need a written list of steps/actions required to complete each task
- Experts need a description or representation of the prototype's interface
- Experts need to have an idea who users are

  - They need to imagine users' thoughts, experiences and their context when interacting with the product

(Wharton et al 1992; Shneiderman et al. 2014)

Cognitive Walkthrough (cont'd)

- The evaluation phase: the experts will ask questions:
  - Are the objects and controls to perform tasks visible? Does the user know this action is a possible choice?
  - Is the location of the identifier (e.g. a label) obvious?
  - Is there a link between the identifier and the action?
    - Sometimes the UI elements are visible but the user doesn't know what will happen if they interact with them
  - Will the user know what actions to perform and their sequence?
  - Is there sufficient feedback so the user knows if the action was completed successfully?
- The interpretation phase:
  - All the information will be gathered and interpreted to identify the easy steps, and difficult steps that require changes

(Wharton et al 1992)

Heuristic Evaluation

- In Heuristic Evaluation, experts use a predefined set of rules to assess the usability of a product interface (a new or an existing system)

- Two well-known Heuristics:

  - **Nielsen's 10 heuristics**

  - **Shneiderman's Eight Golden Rules**

Nielsen's Usability Heuristics

- Nielsen designed heuristic evaluation as a low-cost option for evaluating interfaces.

- He identified a number heuristics (rules of thumb) to assess various interface elements.
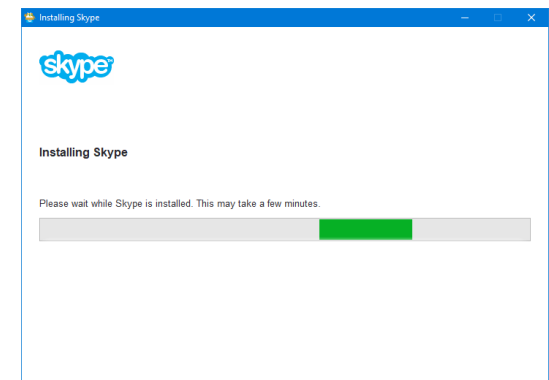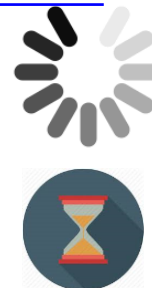
(Nielsen 1995)

Jakob Nielsen
Source: https://www.nngroup.com/articles/author/jakob-nielsen/

Nielsen's 10 Heuristics

Provide feedback
(Nielsen & Molich 1990)

1. **Visibility of system status** (Nielsen 1995)

- Provide users with meaningful and useful feedback about the system and tasks status

- Feedback should be visible, clear, concise and in a timely manner
  - What is going on? What happened?
  - Importance of immediate feedback

- Examples of good practices:
  - When an action is performed, let the user know their interaction was successful
  - Indicate progress or the path in multi-page processes (e.g. breadcrumbs)
  - For longer operations, use a progress bar
 - Indicate when the task is completed

Video https://www.nngroup.com/videos/usability-heuristic-system-status/

FIT5152-Lec4-navigation-menu-g... (17,263K)   ✕

graphologytest.pdf (71K)   ✕

IMG_4959.JPG (2,299K)   ✕

Installing Skype

skype

Installing Skype

Please wait while Skype is installed. This may take a few minutes.

14

Nielsen's 10 Heuristics (cont'd)

> - **Simple and natural dialogue**
> - **Speak the user's language**
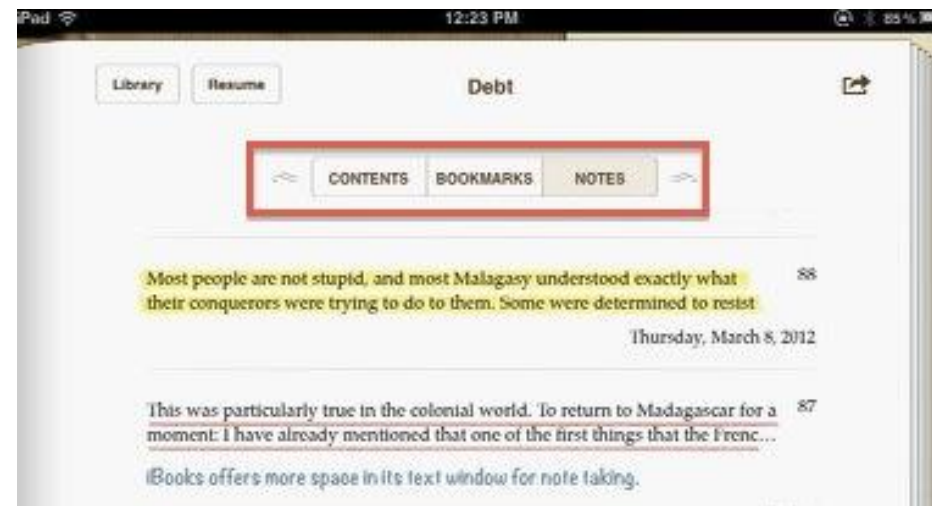>   (Nielsen & Molich 1990)

2. Match between system and the real world (Nielsen 1995)

a. Use of words (language), concepts, and objects familiar to the user

b. Follow real world conventions (skeuomorphic design)

- **Video** https://www.nngroup.com/videos/match-system-real-world/
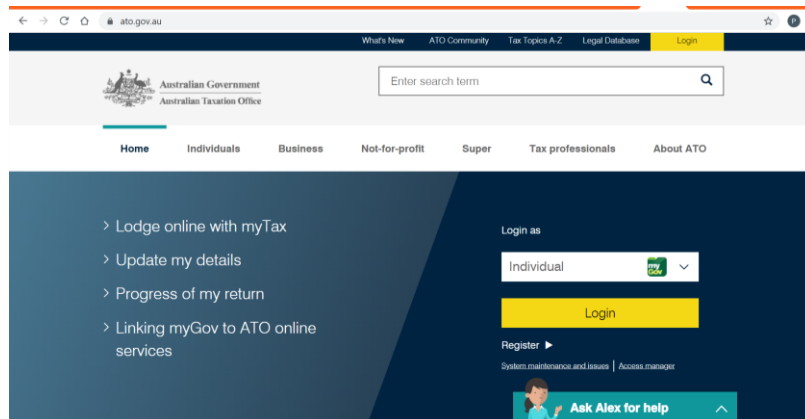
Australian Tax Office (a)

eBooks (b)

15

Nielsen's 10 Heuristics (cont'd)

Provide clearly marked exits (Nielsen & Molich 1990)

3. User control and freedom (Nielsen 1995)

- Users can make mistakes so provide a way to exit from 'unwanted' situations
- Provide options of undo and redo
- Allow users to confirm or cancel actions
- [Video](https://www.nngroup.com/videos/usability-heuristic-user-control-freedom/) https://www.nngroup.com/videos/usability-heuristic-user-control-freedom/
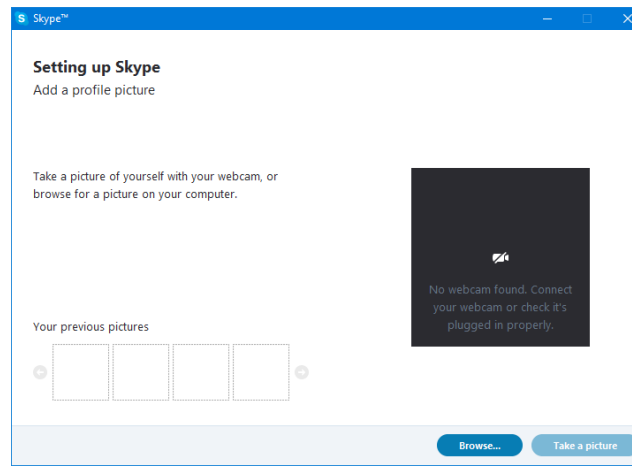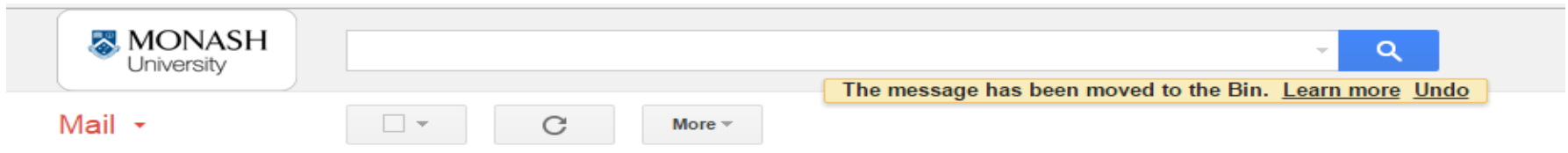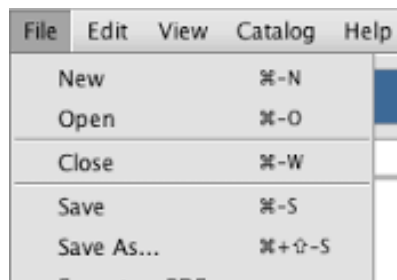
MONASH
University

4. Consistency and standards (Nielsen 1995)

- Follow common standards and conventions.
  - E.g. Mac and Windows use strong consistency to help users learn and use their software and products

- Do not introduce different and unfamiliar words, icons and actions for the same objects.

Standard and common order of
menu options, and familiar words

Familiar icons

File	Edit	View	Catalog	Help
New		⌘-N		
Open		⌘-O		
Close		⌘-W		
Save		⌘-S		
Save As…		⌘+⇧-S		

Nielsen's 10 Heuristics (cont'd)

Prevent errors
(Nielsen & Molich 1990)

5. Error prevention (Nielsen 1995)

Design such that errors are prevented:

- Use constraints
- Use helpful suggestions (auto-suggest or auto-complete features)
- User selected data
- Provide examples of data entry format
- Use of useful defaults, particularly for doing repetitive actions
- Use validation of data entry

Nielsen's 10 Heuristics (cont'd)

6. Recognition rather than recall (Nielsen 1995)

- Recall requires the user to retrieve information from the memory

- Recall can be error-prone and difficult

- Recognition reduces memory load by making actions and options visible

- Menus are the most classic example of recognition
  - E.g. to underline text, you look at the menu and recognise the option

- Providing recognition in interfaces gives the user extra help in remembering information
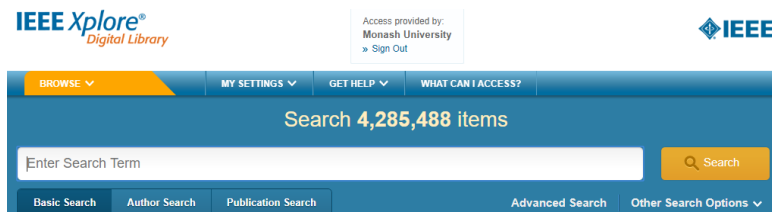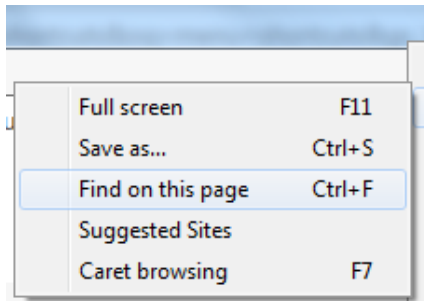
- Promote recognition over recall in your UI design

Monash Gmail
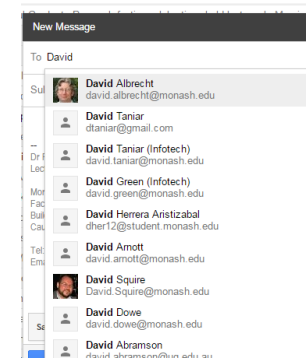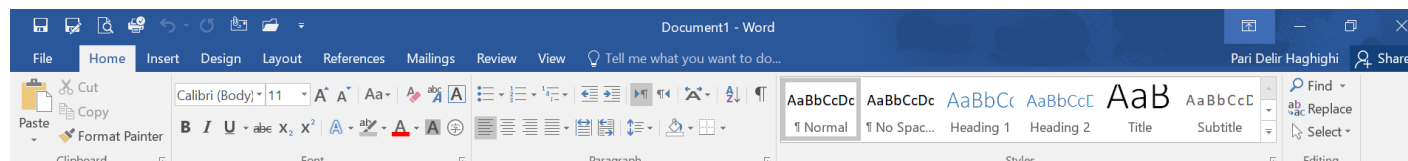
Nielsen's 10 Heuristics (cont'd)

7. **Flexibility and efficiency of use** (Nielsen 1995)

- The system should support both novice and experienced users.
  - E.g. providing keyboard shortcuts
  - E.g. advanced search options
  - **Video** https://www.nngroup.com/videos/flexibility-efficiency-use/

Full screen	F11
Save as...	Ctrl+S
Find on this page	Ctrl+F
Suggested Sites	
Caret browsing	F7

IEEE Xplore® Digital Library

Access provided by:
Monash University
» Sign Out

◆IEEE

BROWSE ˅   MY SETTINGS ˅   GET HELP ˅   WHAT CAN I ACCESS?

Search **4,285,488** items

Enter Search Term   🔍 Search

Basic Search   Author Search   Publication Search   Advanced Search   Other Search Options ˅

Advanced Search Options

| Advanced Keyword/Phrases | Command Search | Citation Search | Preferences |

ENTER KEYWORDS OR PHRASES, SELECT FIELDS, AND SELECT OPERATORS
Note: Refresh page to reflect updated preferences.

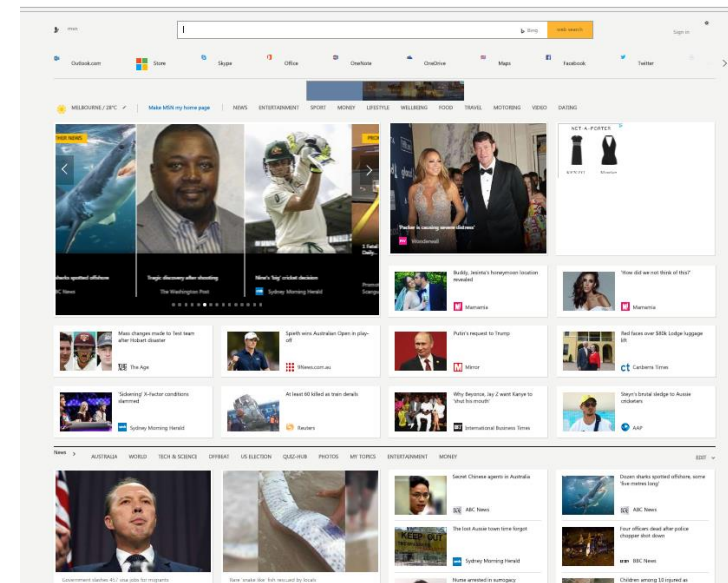Search : ● Metadata Only ○ Full Text & Metadata

	in Metadata Only ▼
AND ▼	in Metadata Only ▼
AND ▼	in Metadata Only ▼

+ Add New Line   Reset All   **SEARCH**

8. Aesthetic and minimalist design (Nielsen 1995)

- Avoid including information that is irrelevant or might be rarely used
- Avoid overloading and cluttering

Google
vs MSN

Nielsen's 10 Heuristics (cont'd)

Good error messages
(Nielsen & Molich 1990)

9. Help users recognize, diagnose, and recover from errors
(Nielsen 1995)

- Error messages clearly specify what is the problem (in a language that any user can understand), and provide a useful option or solution.
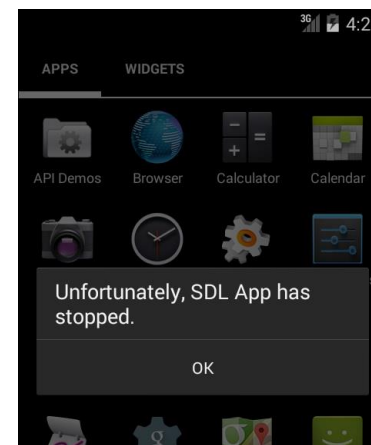

The blue screen of death (BSOD)

Android emulator

??

22
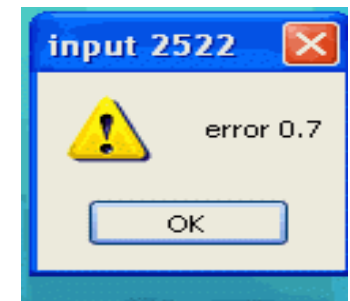
Nielsen's 10 Heuristics (cont'd)

10. Help and documentation (Nielsen 1995)

- Although the system can be used without documentation, it is recommended to provide it.

- Help and documentation should provide effective and efficient search options to find information easy and fast according to the user task and keyword.
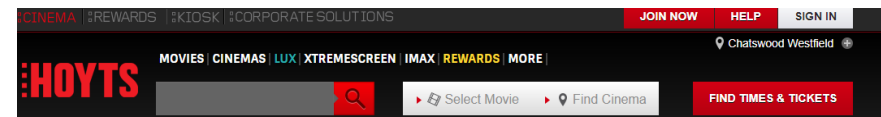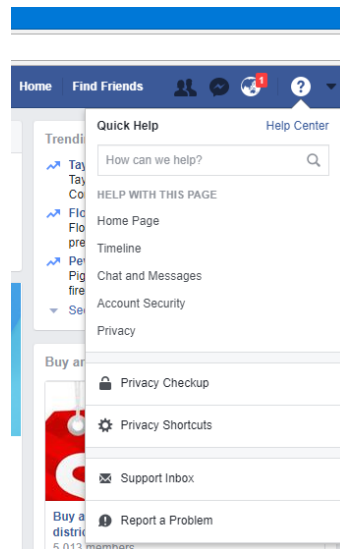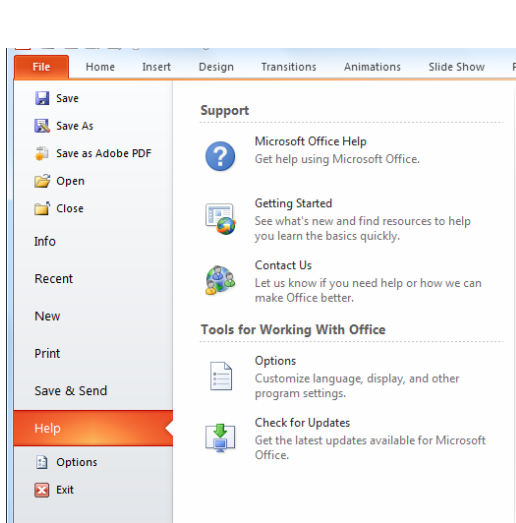
The Eight Golden Rules

1. Strive for consistency
2. Cater to universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions (undo)
7. Support internal 'locus of control'
8. Reduce short term memory

Shneiderman et al (2014)

Ben Shneiderman
Source Twitter

24

The Eight Golden Rules in Detail

1. Strive for consistency
   – Consistency of terminology, colour, layout, fonts, or order of actions
2. *Cater to universal usability
   – Design for diverse users, and consider different levels of experience, age difference, disabilities, and cultural differences
3. Offer informative feedback
   – For more frequent and minor actions, provide low level of feedback
   – For infrequent and major actions, provide detailed feedback
4. Design dialogs to yield closure
   – This is about organising a series of actions into groups (the beginning, middle and end), and providing right feedback according to each stage, e.g. a confirmation at the end

*2. Enable frequent users to use shortcuts

Shneiderman et al (2014)

5. Prevent errors (offer simple error handling)

- Design the system such that errors can be avoided
- Detect errors and handle the error in a simple way

6. Permit easy reversal of actions (undo)

7. Support internal 'locus of control'

- Users should feel they are in charge of the interface
- Provide them with more flexibility and options
- Do not surprise them with unexpected results and behaviour

8. Reduce short term memory

- Similar to recognition rather than recall

Shneiderman et al (2014)

Compare Heuristic Evaluation Rules

Shneiderman 8 Golden Rules

1. Strive for consistency
2. Cater to universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions (undo)
7. Support internal 'locus of control'
8. Reduce short term memory

Nielsen Heuristics

1. Visibility of system status
2. Match between system and real world
3. User control and freedom
4. Consistency and standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility & efficiency of use
8. Aesthetic & minimalist design
9. Help users recognize, diagnose, recover from errors
10. Help and documentation

Heuristic Evaluation Steps

- The evaluation team members separately inspect the interface according to the evaluation criteria to determine if it conforms or violates any

- Evaluators usually go through the interface twice
  - The first pass to understand the interaction and task flow and the system scope
  - They identify the main tasks
  - The second pass to focus on specific interface elements

- Evaluators then gather and consolidate their results

- The problems will be identified and ranked on a scale of 0- 4

The Output of Heuristic Evaluation

- The output will be a list of usability problems in the interface
  - The problem definition
  - Explain which heuristic was violated or not fully followed
  - Evidence of the above
  - Severity Ratings (0-4)
  - Evaluators provide recommendations to revise and improve the interface

  The information above should be specific and provided for each usability problem separately

The Severity of the Usability Problems

- 3 factors help to determine the severity of a usability problem:
  - **Frequency**
    - "how many users will be affected by the problem?" (Hertzum 2006, pg 3)
    - Is this problem common or rare?
      - It can be calculated by the number of evaluators in a test who experience a problem divided by the total number of evaluators
  - **Impact**
    - The impact on the task and the user (low or high impact)
    - Is it easy or difficult to deal with the problem?
    - Rating the impact is less objective
  - **Persistence**
    - "how many times will a user experience the problem?" (Hertzum 2006, pg 3)
    - Is the problem recurring or one time issue?

https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/
http://uxpajournal.org/the-relationship-between-problem-frequency-and-problem-severity-in-usability-evaluations/

Nielsen's Severity Rating

0 = I don't agree that this is a usability problem at all

1 = Cosmetic problem only: need not be fixed unless extra time is available

2 = Minor usability problem: fixing this should be given low priority

3 = Major usability problem: important to fix, so should be given high priority

4 = Usability catastrophe: imperative to fix this before product can be released

https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/

The Heuristic Evaluation Report

1. A summary table

#	Problem	Heuristic Rule	Evidence	Severity Rating	Recommendation
1	Not all the fields have tooltips in the user data entry form	#4 and #10	Figure 1: the entry field where the tooltip was not used	2	The tooltip should be added
…	…	…	…	…	…

2. An explanation of each part

- Problem definition:
  - Describe the problem, which rule it violates and why
- Evidence
  - Provide screenshot/s and use arrows and boxes to highlight the problems
- Severity rating
  - In terms of frequency, impact and persistence
- Recommendation
  - Describe how to fix the problem

Issues with Heuristic Evaluation

- It can be difficult to find experts
- Experts can have biases
- Important problems may get missed
- Different evaluators can find different problems
- Not always possible to identify the problems real users would encounter
- Many different heuristics now available which ones are the 'best'?
  - Heuristics can be platform dependent

Evaluation Types

- When evaluators are users:
  - Usability Testing
  - Field studies
  - Thinking aloud

Usability Testing

- Usability testing was introduced in early 1980s, realising how important user needs are in the interface design.

- Easier and cheaper to fix problems during the design cycle rather than after the system is implemented.

- Usability testing aims to evaluate whether a product meets user needs and its intended purpose (the usability aspects: effectiveness, efficiency, and satisfaction).

Effectiveness: the accuracy and completeness with which users achieve goals within a context
Efficiency: based on the resources expended (e.g. time and effort) to complete a task
Satisfaction: the level of the comfort and acceptability of the system to its users

Designing a Usability Test with Real Users

1. What is the purpose of this test? What are the user goals to achieve? What do you plan to test/measure?

2. Who are your users? How will you recruit them? How many users/participants?

3. The testing environment and setting? Equipment? (e.g. mobile app testing requires mobile phones)

4. Identify the list of tasks to be performed by users related to the goals. What do users need to do? The task selection, order, and design

5. The structure of the test - What is the procedure?

6. Survey instrument: how will you collect the data/feedback? Questionnaire or interviews or both?

7. Analysis methods and statistics (Qualitative or Quantitative)

8. Ethics approval

9. A Consent Form and an Explanatory Statement should be obtained

   • Ask participants to read and sign them before starting the test

Basic Principles

Real **users**

– Users should be those currently using the product or who will later use it.

– If the user group is unknown, or very broad e.g. a website then must select representative users

– If participants are more experienced than actual users, they may miss problems.

– Participation in usability tests are voluntary.

Real **tasks**

– Select tasks that users would likely do.

– Tasks can be broken into subtasks.

– Decide on tasks order and description.

– For complex systems, may only test some of the tasks, and select users who are most likely to do those tasks.

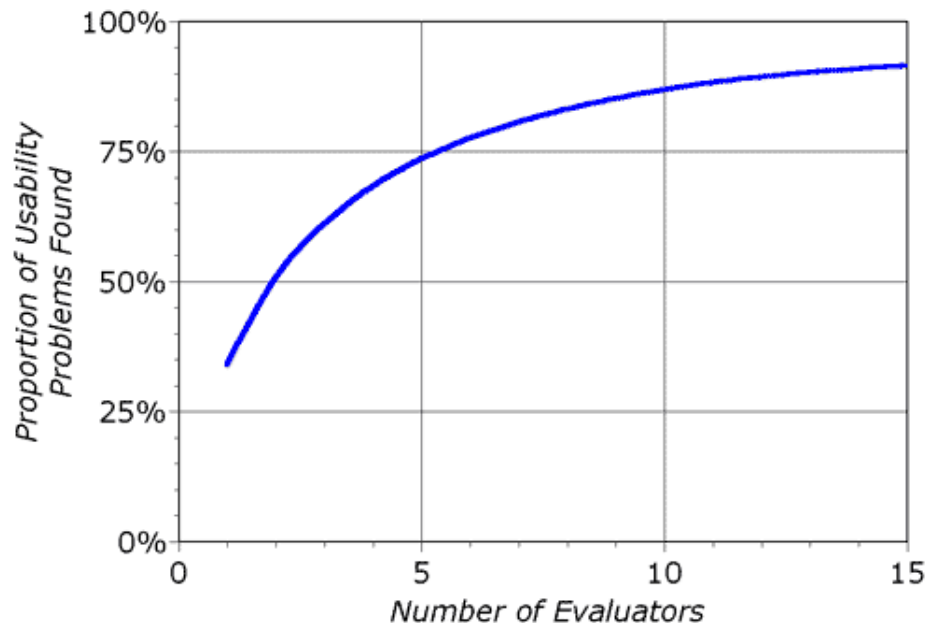# Who are your users?

Video of Usability test with a child

Number of Evaluators

- For qualitative analysis: 3–5 users
- If usability metrics are used: 20 users are recommended

(Nielsen , 2000)

https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/
https://www.nngroup.com/articles/how-many-test-users/
https://www.nngroup.com/articles/quantitative-studies-how-many-users/

What to Measure?

- Measuring usability based on a set of usability criteria
- Many different set of criteria proposed by researchers
  - Need to select the right criteria for the test
- Most of these criteria focus on 3 main aspects:
  - Effectiveness:
    - Success rate (task completion rate)
    - Number of errors per task
  - Efficiency:
    - How long each task took, and the overall time
  - Subjective measures (more qualitative analysis):
    - People's satisfaction level (task level), perceptions, opinions and judgments
      - Common to use a Likert scale
      - Open-ended questions

Examples of What to Measure

Time

- To finish a task
- Spent navigating menus
- Spent in 'help'
- Trying to find information
- Recovering from errors

Number of

- clicks users make
- pages they browse
- screens looked at online help
- wrong menu choices
- wrong function keys chosen
- repeated errors

Eye Tracking

- Measuring where the eye is focused on the page
- How long, where, navigation?

Tobii Pro Glasses 2

Number of

- percentage of tasks that users complete correctly
- Ratio of successes to failures

Video

- Observations

Measuring Usability: Examples

Subjective measurements:

- Ratings of:
  - *Ease of learning*
  - *Ease of using and doing*
  - *Ease of installation*
  - *Helpfulness of online help*
  - *Ease of understanding information*

- Preferences and reasons for preferences:
  - *Over the previous version*

- Predictions of behaviour and reasons for behaviour:
  - *Would you buy the product?*
  - *Would you use this website again?*

- Spontaneous comments:
  - *"I am totally lost here"*
  - *"that was easy"*
  - *"I don't understand this message"*

Examples of usability questions for a Website

Open-ended questions:

- What was the best feature or part of the website?
- What was the worst feature or part of the website ?
- If you would not use this website again please explain why not.

Yes/No questions

- Did you get lost at any stage during your exploration of the site?
- Were there any aspects of the website that caused confusion or slowed down your retrieval of information?
- Was there anything else you wanted to know but could not find out from the site?

Examples of Likert Scale questions 1-5
(1 disagree strongly – 5 agree strongly)

Strongly disagree	○	Disagree	○	Neutral	○	Agree	○	Strongly agree	○

- The size of the text was easy to read
- The text was displayed in a way that was easy to read
- It was easy to navigate through the site
- The site was easy to use
- Generally there was too much information on the screen
- The language used was easy to understand
- The design of the interface was appealing
- The graphics on the site were appealing
- Different parts of the interface such as the icons were consistent
- All the information I required to complete the task was on the website
- The number of steps required to get to the information I wanted was acceptable
- I understood the terminology used on the website

Task Description Example

Task #	Task Description
#1	Login Click on Login button to continue.
#2	Data Entry Choose Daily Data Entry option from the menu.

#3	Enter your daily record
	
	   1. Blood Pressure
	
	      Choose Manual Data Entry option to enter your blood pressure values.
	
	      Enter 128 for Systolic field
	
	      Enter 84 for Diastolic field
	
	      Enter 70 for Pulse field
	
	   2. Press NEXT button and select CAPD option under Type of Dialysis field.
	
	   3. Press SUBMIT button.

#4	Check your daily report
	
	   4.1 Select Progress Reports option from the menu list
	
	   4.2 Select Daily Report option from the list.
	
	   4.3 Click on Back button.

#5	Check your weekly report
	
	   5.1 Select Weekly Report option from the list.
	
	   5.3 Tap on calendar icon and select start date as 5-Sep-2016
	
	   5.4 Tap on calendar icon and select end date as 15-Sep-2016
	
	   5.5 Tap on Get Progress Report button

Tap on Back button of the phone to complete this part.

Questionnaire Example

|	Statements	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
B1	Were you able to complete the set of tasks?					
B2	Were you confused while filling out details on screen?					
B3	Was the application easy to learn?					
B4	Was it easy to enter the text/numeric value manually?					
B5	Was text readable?					
B6	Was text easy to read with large font size?					
B7	Was colour used in interface made text understandable and easy to read?					
B8	Was colour used in interface attractive?					
B9	Was navigational structure simple by using NEXT button?					

Data Analysis Examples

Quantitative approaches:

- Descriptive statistics: e.g. mean, frequencies (or relative frequency %), graphs, or cross/pivot tables
- Statistical tests

Qualitative approaches:

- Coding and Theming

Percent of Users Who Started and Completed Each Step	
Step 1	91%
Step 2	85%
Step 3	53%
Step 4	45%
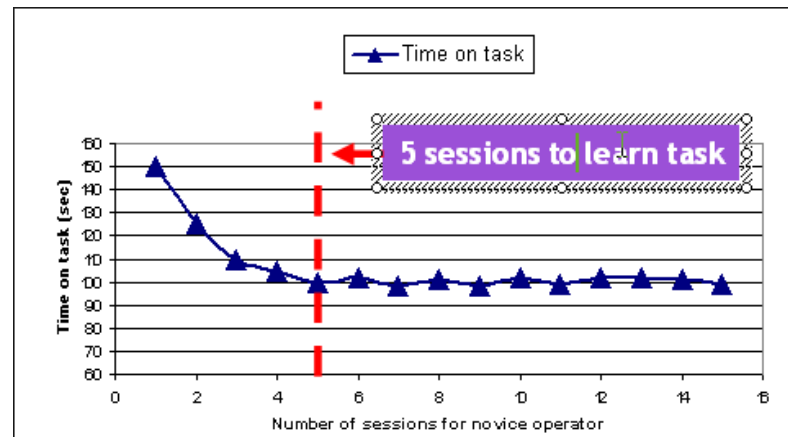
Explanatory Statement - Example

You are invited to take part in this study. Please read this Explanatory Statement in full before deciding whether or not to participate in this research. If you would like further information regarding any aspect of this project, you are encouraged to contact the researchers via the phone numbers or email addresses listed above.

What does the research involve?

This study aims to collect feedback to evaluate our mobile diary application for patients with chronic kidney disease (CKD) on peritoneal dialysis (PD). The mobile app is developed to facilitate the tasks of entering and managing daily records and aims to improve the user experience by adjusting the mobile app' user interface according to the context of the user such as age or digital literacy skills.

Why were you chosen for this research?

You would have been referred to us by nurses and doctor/s who are involved in your dialysis treatment.

Consenting to participate in the project and withdrawing from the research

Being in this study is voluntary and you are under no obligation to consent to participation. However, if you do consent to participate, you may withdraw from further participation at any stage but you will not be able to withdraw data that has been collected.

Possible benefits and risks to participants

There will be no direct benefits for participants. However, findings from this research will provide a better understanding of mobile application interfaces and help to improve the existing user interface guidelines for mobile app development.

We believe that there will be no stress, inconvenience or discomfort to any of the participants.

Confidentiality

All information collected in this study will be de-identified. Collected data will be used for scientific publications while maintaining the anonymity of the data sources. The information collected on the consent form will not be published, or released in any way.

Storage of data

The data will be stored in accordance with Monash University Regulations, kept in a secured filing cabinet. A report of study or data may be submitted for publication, but individual participants will not be identifiable in such a report.

Use of data for other purposes

The data collected may be used for other purposes in a completely de-identified form, subject to the approval of the Chief Investigator, where ethics approval has been granted.

Results

Participants may opt to be informed of the research findings by contacting us by email:

Complaints

Should you have any concerns or complaints about the conduct of the project, you are welcome to contact the Executive Officer, Monash University Human Research Ethics (MUHREC):

> Executive Officer
> Monash University Human Research Ethics Committee (MUHREC)
> Room 111, Chancellery Building E,
> 24 Sports Walk, Clayton Campus
> Research Office
> Monash University, VIC – 3800
>
> Tel: +61 3 9905 2052 Email: muhrec@monash.edu Fax: +61 3 9905 3831

Thank you,

Consent Form - Example

Project title:

Investigators names:

I have been asked to take part in the Monash University research project specified above. I have read and understood the Explanatory Statement and I hereby consent to participate in this project.

I consent to the following:	Yes	No
I am willing to be interviewed by the researcher	☐	☐
I agree to allow the interview to be audio-taped	☐	☐
I agree to complete the questionnaire	☐	☐
I give my permission for my data to be used for future research	☐	☐

And
I understand that my participation is voluntary, that I can choose not to participate in part or all of the project, and that I can withdraw at any stage of the project without being penalised or disadvantaged in any way,

And
I understand that data from the questionnaire and interview will be kept in secure storage and accessible to the research team. I understand that data collected will be kept in secure storage and is accessible to the research team. I also understand that the data will only be used for future research if I consent to it.

Name of Participant: _____

|

Signature: _____Date_____

Usability Testing in Laboratories

- The usability lab includes a testing room and an observation room.

- Controlled environmental settings where users are observed.

- Recording or monitoring users' performance doing typical tasks.

- Data is stored and analysed.

- Measuring performance times to identify & explain errors.

Field Studies

- It takes the usability testing and research to the participants and their environment.

- It involves observing users in their natural setting.

- Increases understanding of what users do and the impact of technology.

- It can involve interviews.

- Most important part of being there is to observe, ask questions and record what users are doing and saying.

- Research shows that field testing works well with older adults.

Source http://uxpamagazine.org/field_studies_older_adults/

Usability Testing In-The-Wild

- The shift from usability testing in the labs to wild studies
- This type of testing involves recording and observing users over a period of time
- It is about testing prototypes of new systems, devices or services that are used in people's every day life, and how they are adopted (rather than their usability)
- E.g. usability testing of mobile devices, smartphones, smartwatches and wearables

(Rogers 2011)

Usability Testing In-The-Wild (cont'd)

- Challenges:
  - How long to evaluate the users?
  - There is a lack of a facilitator during the test
  - In the wild, the test can be affected by a number of 'factors and interdependencies' (not a controlled environment like a lab)
  - Users are 'unpredictable' in the wild, can be 'interrupted or interrupt their own activities'

    (Rogers 2011, pg 58)

  - Long term wild studies can be expensive
  - Testing over a long period of time can be uncomfortable for users
  - Ethical issues possible

    (Chamberlain et al. 2012)

Thinking Aloud

- An effective method that can be used during usability testing
- Users will think loud when performing tasks and interacting with the system
- Testers will listen for the clues
  - E.g. *Hmm … how I go back? … maybe on the top left… I give it a try...*
- Questions and comments at the end
- Verbalising can create additional cognitive load for users

Shneiderman et al (2014)

Video of testing Tripadvisor.com and thinking aloud

Summary

- Expert reviews and heuristic evaluation can be very useful in revealing usability issues and improving the system but it cannot be as effective as usability testing

- Usability testing costs money and takes time but is crucial for successful interfaces

*"Designers can become so entranced with their creations that they may fail to evaluate them adequately. Experienced designers have attained the wisdom and humility to know that extensive testing is a necessity."

(Shneiderman and Plaisant 2005 pg 140).

References

- Wharton, C., Bradford, J., Jeffries, R., & Franzke, M. (1992, June). Applying cognitive walkthroughs to more complex user interfaces: experiences, issues, and recommendations. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 381-388). ACM.
- Tomayess Issa, Pedro Isaias (2015) Sustainable Design HCI, Usability and Environmental Concerns, Springer
- Shneiderman, Plasiant, Cohen, Jacobs (2014) Designing the User Interface: Strategies for Effective Human-Computer Interaction, Pearson (Chapter 4)
- Nielsen, J. (1992). Finding Usability Problems Through Heuristic Evaluation. *Proceedings of the SIGCHI conference on Human factors in computing systems CHI ' 92*, Monterey, ACM.
- Nielsen, J. (1989, September). Usability engineering at a discount. In *Proceedings of the third international conference on human-computer interaction on Designing and using human-computer interfaces and knowledge based systems (2nd ed.)* (pp. 394-401). Elsevier Science Inc.
- Chamberlain, A; Crabtree, A; Rodden, T; Jones, M; Rogers, Y; (2012) Research in the wild: Understanding 'in the wild' approaches to design and development. In: (pp. pp. 795-796).
- Hertzum, M. (2006). Problem prioritization in usability evaluation: From severity assessments toward impact on design. International Journal of Human-Computer Interaction, 21(2), 125-146.
- Usability testing – Success Rate https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric/
- User Experience Direct (UX Direct) FAQ: How to conduct Heuristic Evaluation, Oracle http://www.oracle.com/webfolder/ux/applications/uxd/assets/faq/how-to-conduct-heuristic-evaluation.pdf
- Xu Sun and Andrew May, A Comparison of Field-Based and Lab-Based Experiments to Evaluate User Experience of Personalised Mobile Devices https://www.hindawi.com/journals/ahci/2013/619767/