



The
University
Of
Sheffield.

COM2002/3002/6222

Human-Centred Systems Design

Dr Maria-Cruz Villa-Uriol

Bibliography



Evaluation in HCI

- In the upcoming 4 Tuesday lectures we will cover:

Measuring Usability



Qualitative Research Methods



Experimental Design



Quantitative Research Methods

Week 8 →

Assignment
hand-out

Week 9

Week 10

Week 11 →

Assignment
hand-in
via MOLE

Assessment

- Hand-out (via MOLE):
Tuesday 14th November (week 8)
- Hand-in (via MOLE):
Monday 11th December (week 12)
- 5 questions, maximum 100 words (essay-like).
- 20% towards the final mark



The
University
Of
Sheffield.

COM2002/3002/6222
Human-Centred Systems Design

Measuring Usability

Dr Maria-Cruz Villa-Uriol

Today...

- We will cover:
 - Usability
 - Usability Engineering
 - Usability Testing
 - Evaluation
 - Goals
 - Types

Usability

True or false?

- **Usability** is an fabrication of designers meant to complicate software developers lives.
- **Usability** is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.
- **Usability testing** is often known as “user research”.
- **Usability testing** is a way to see how easy to use something is by **testing** it with real users. Users are asked to complete tasks, typically while they are being observed by a researcher, to see where they encounter problems and experience confusion.
- **Usability testing** is useless.



Heinz case study

efficient

effective

easy to learn



safe to operate



high utility

Usability

- ‘**Usability**’ is a central pursuit of Human Computer Interaction
- The original definition (*Shackel, 1990*) is that systems should ...
 - be easy to use
 - be easy to learn
 - be flexible
 - engender a good attitude in people
- As the variety of **People**, **Activities**, **Contexts** and **Technologies** of interactive system design has increased, so this definition hides important issues such as ...
 - accessibility
 - sustainability
- The ‘**goals of usability**’ are now primarily seen as concerned with system ...
 - efficiency
 - effectiveness

Shackel, B. (1990). Human factors and usability. In J. Preece & L. Keller (Eds.), *Human-Computer Interaction: Selected Readings*. Hemel Hempstead: Prentice Hall.



Usability

- ‘**Accessibility**’
 - removing the barriers that would otherwise exclude some people from using the system at all
- ‘**Usability**’
 - the quality of the interaction in terms of parameters such as ...
 - time taken to perform tasks
 - number of errors made
 - time to become a competent user
 - a system must be accessible before it is usable
 - a system may be assessed as highly usable according to some usability evaluation criteria, but may still fail to be adopted or to satisfy people
- ‘**Acceptability**’
 - fitness for purpose in the context of use
 - also covers personal preferences that contribute to users ‘taking to’ or ‘forming a liking for’ an artifact, or not



Usability *vs.* Acceptability

- ‘**Acceptability**’ is about fitting technologies into people’s lives, e.g. ...
 - some railway trains have ‘quiet’ carriages where it is unacceptable to use mobile phones
 - cinemas remind people to turn their phones off before the film starts
 - Apple’s iMac computer was the first computer designed to look good in a living room
 - a computer playing loud music would generally be considered to be unacceptable in an office environment
- Acceptability can *only* be understood in the context of use
- Usability can be evaluated in a laboratory, but acceptability cannot



Google vs. Yahoo


[My Yahoo!](#) [My Mail](#)
[Web](#) | [Images](#) | [Video](#) | [Audio](#) | [Directory](#) | [Local](#) | [News](#) | [Shopping](#)
Search:
[Yahoo! Answers: What is the difference between a road and a street?](#)
[Web Search](#)

2006

- [Autos](#)
- [Entertainment](#)
- [Finance](#)
- [Games](#)
- [GeoCities](#)
- [Groups](#)
- [Horoscopes](#)
- [HotJobs](#)
- [Kids](#)
- [Maps](#)
- [Personals](#)
- [Photos](#)
- [Real Estate](#)
- [Shopping](#)
- [Sports](#)
- [Travel](#)
- [Yellow Pages](#)
- [**All Yahoo! Services**](#)

Featured [Entertainment](#) [Sports](#) [Money](#)



Who will your team pick?

Y! Sports' Charles Robinson breaks down Saturday's first round. [more](#)

 Cris Carter: Who I'd take No.1...
 Dan Wetzel: Young and priceless
 More Y! Sports NFL Draft analysis

 NFL Draft: Who will your team take?

 VIDEO: 'Da Vinci' judge hides code in ruling

 Tips for fighting allergies

 'Idol' viewers boot Pickler, leaving five

In the News [World](#) [Video](#)

- Bush visits New Orleans as FEMA criticized
- Congress struggles to act on gas prices
- Rice says U.N. must act over Iran nukes
- Sago air equipment has history of problems
- Prominent Belarus opposition leader jailed
- Colombians seek refuge in squatter villages
- Ten states sue EPA over global warming

[» More News](#)

Markets: Dow: **+0.2%** Nasdaq: **+0.5%**

Stock Quotes: [Go](#)

Marketplace



\$300? \$400? \$500?
How much could you save with GEICO? Get an instant quote at [GEICO.com](#) now.

Get your degree online in less than 1 year - A.S., B.S., M.S. - Start

Check your mail status: [Sign In](#) Free mail: [Sign Up](#)

 Mail	 Messenger	 Music
 Weather	 Local	 Movies

Yahoo! Photos

Bring your photos to life. Share memories with friends and family.

- Free, unlimited storage
- Easy uploading
- Available in over ten countries

[» Start sharing](#)



Yahoo! Travel
Plan a Summer Trip

Top 10 Beaches
National Parks Guide
Summer Road Trips
Romantic Getaways
More Travel

Special Offer

BLUE SKY
AMERICAN EXPRESS

Book any trip.
Anytime. Anywhere.

Yahoo! Pulse

Most Popular Sedans

1	Honda Civic	Read user reviews
2	BMW 3 Series	Read user reviews

2006

[philipp.lenssen@gmail.com](#) | [Personalized Home](#) | [Search History](#) | [My Account](#) | [Sign out](#)



Web [Images](#) [Groups](#) [News](#) [Froogle](#) [Maps](#) [more »](#)

[Advanced Search](#)
[Preferences](#)
[Language Tools](#)

[Google Search](#) [I'm Feeling Lucky](#)

[Advertising Programs](#) - [Business Solutions](#) - [About Google](#) - [Go to Google Deutschland](#)

©2006 Google

Google interface evolution:
<http://blogoscoped.com/archive/2006-04-21-n63.html>

Usability

- A system with a high degree of usability will ...
 - be **efficient** (*in that people will be able to do things using an appropriate amount of effort*)
 - be **effective** (*in that it contains the appropriate functions and information content, organized in an appropriate manner*)
 - be **easy to learn** (*in how to do things and remember how to do them after a while*)
 - be **safe** to operate (*in the variety of contexts in which it will be used*)
 - have **high utility** (*in that it does the things that people want to get done*)



THE IMPORTANCE OF **USABILITY**



LEARNABILITY

How quickly can you learn to use your software?



EFFICIENCY

How quickly can you perform your tasks?



SATISFACTION

How much do you enjoy using your software?



MEMORABILITY

How quickly can you remember how to use your software?



ERRORS

How often do you encounter errors in your software?

Usability Engineering

- The ultimate test of usability is based on the measurement of user experience
- ‘**Usability engineering**’ demands that specific usability measures be made explicit as *requirements*
- Usability specification ...
 - usability attribute/principle
 - measuring concept
 - measuring method
 - now level / worst case / planned level / best case
- Problems ...
 - usability specification requires level of detail that may not be possible early in design
 - satisfying a usability specification does not necessarily satisfy usability

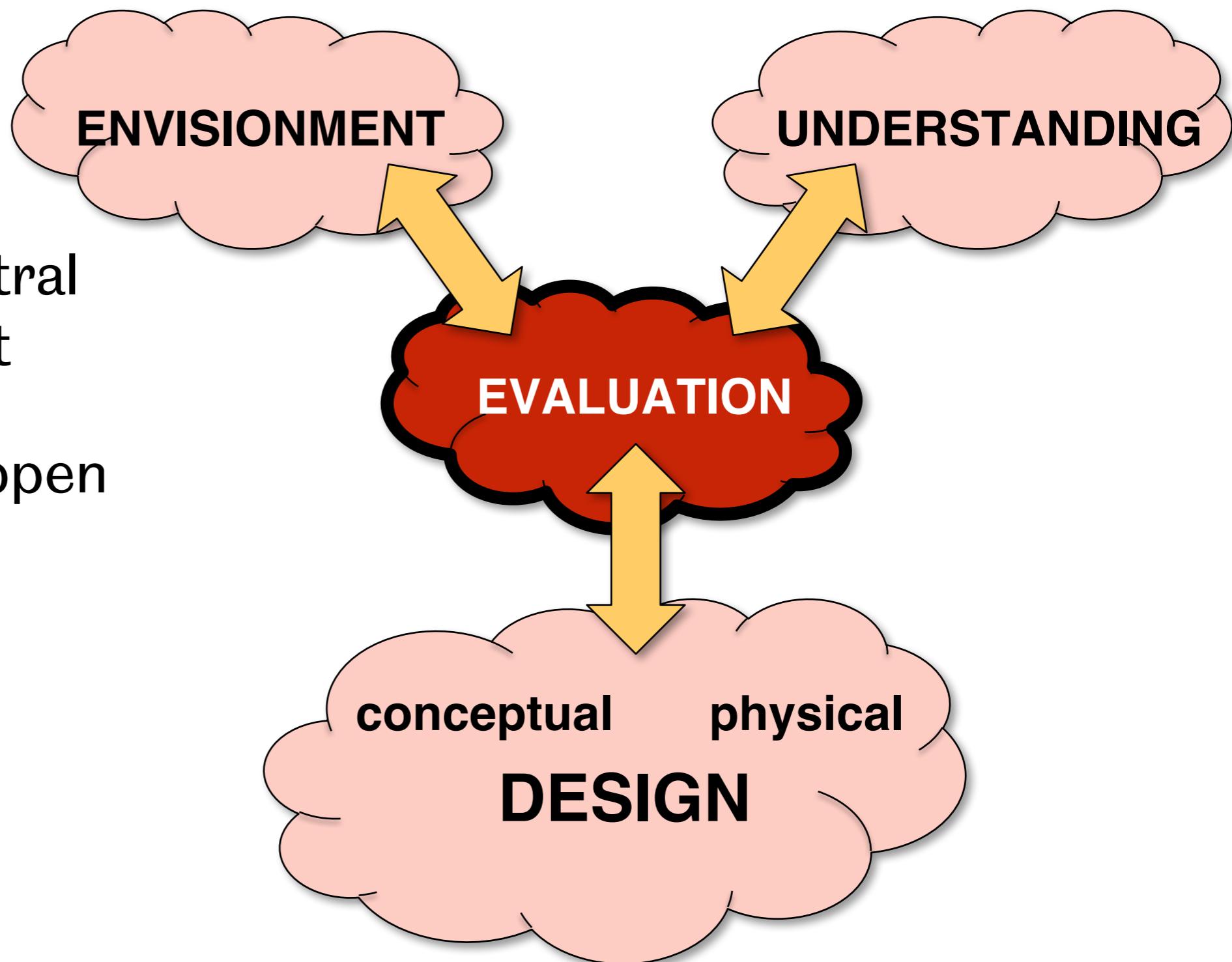
Usability Engineering

Usability Specification for a DVD Recorder

Attribute:	<i>Backward recoverability</i>
Measuring concept:	<i>Undo an erroneous programming sequence</i>
Measuring method:	<i>Number of explicit user actions to undo current program</i>
Now level:	<i>No current product allows such an undo</i>
Worst case:	<i>As many actions as it takes to program-in mistake</i>
Planned level:	<i>A maximum of two explicit user actions</i>
Best case:	<i>One explicit cancel action</i>

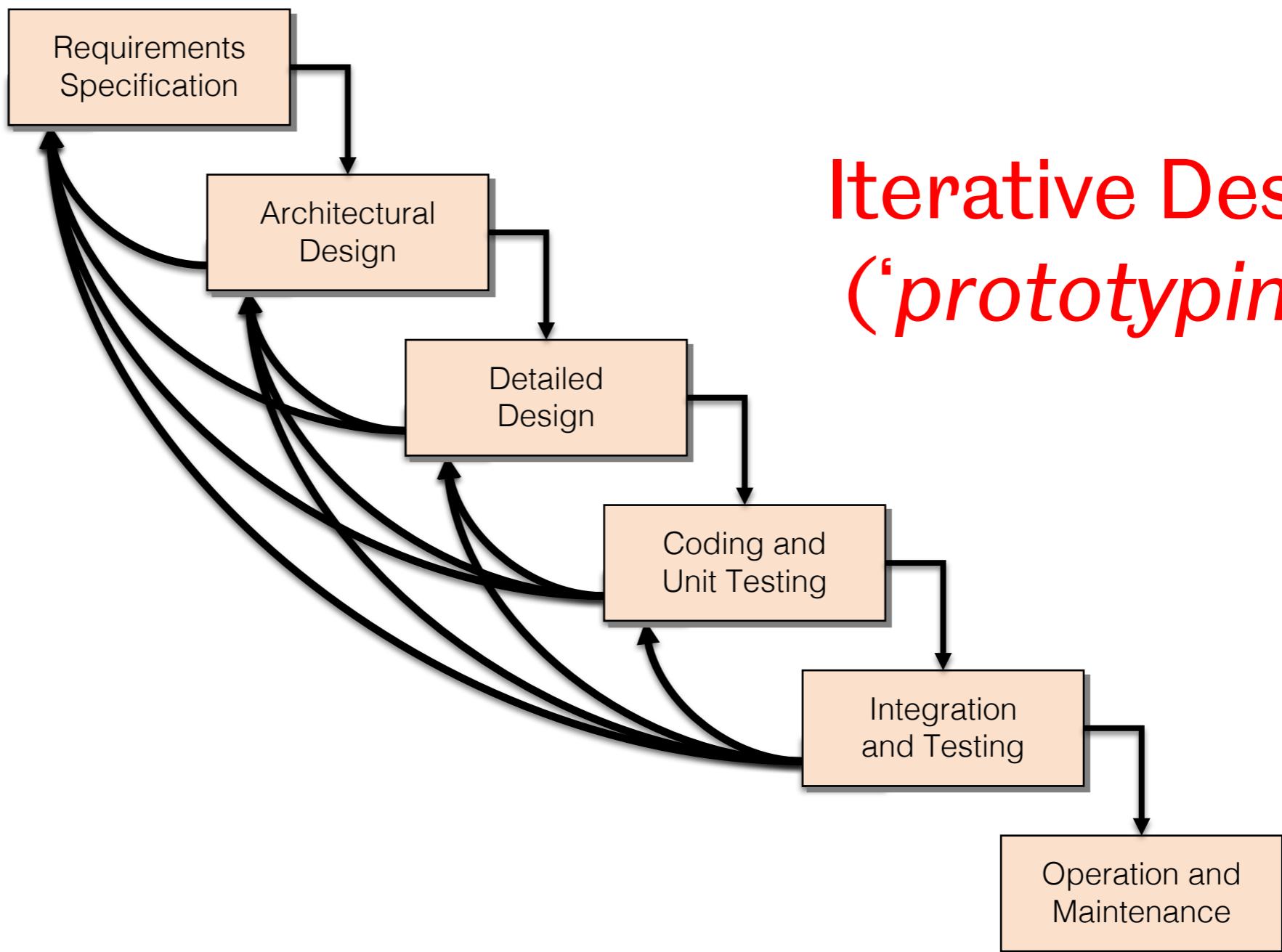
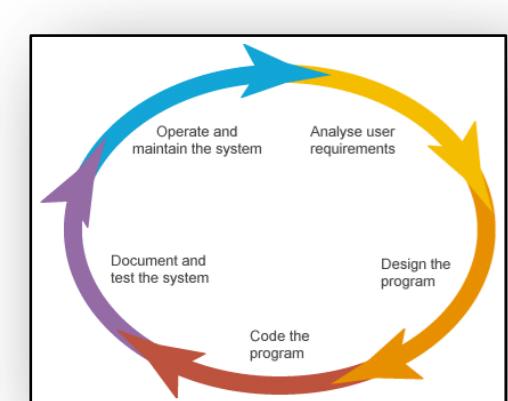
The Design Process

- Evaluation is central
- Process can start at any point
- Activities can happen in any order



The ‘Life Cycle’ for Interactive Systems

**Iterative Design
(‘prototyping’)**

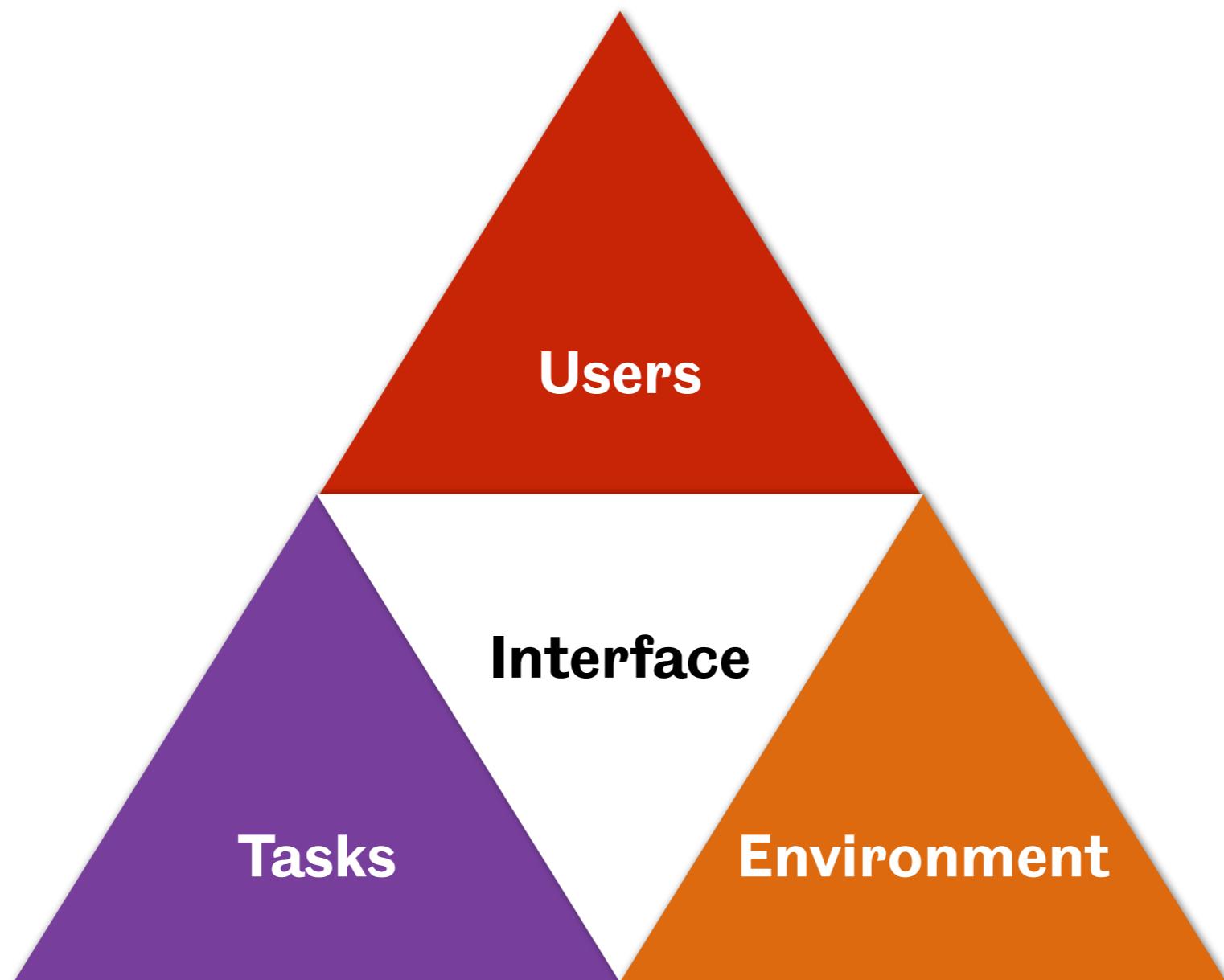


Usability Testing

- **Representative** users attempting **representative** tasks in **representative** environments

Usability Testing

- Representative **users** attempting representative **tasks** in representative **environments**



Usability Testing

- Representative users attempting representative tasks in representative environments
- This includes testing:
 - Paper prototypes
 - Screen mock-ups
 - Prototypes controlled via Wizard of Oz
 - Partially-functional software
 - Fully-working versions of software
 - ...

Usability Testing - Aim

To improve the quality of an interface
by finding flaws to fix it

And AS SOON AS POSSIBLE, please!

Q&A

Is “usability testing” research?

Usability Testing vs. Traditional Research Methods

- Similarities in terms of methods
 - Measures of task performance and time performance
 - Surveys of user satisfaction
 - Observation techniques from ethnography
- Differences as end goals are typically different
 - Usability testing is an industry approach
 - Practical emphasis vs theory
 - Concern with building successful product in shortest time, with fewest resources, while minimizing risks

International Organisation for Standardization (ISO)

- ISO provides requirements, specifications, guidelines or characteristics to consistently ensure that materials, products, processes and services are fit for their purpose
- The British Standards Institution (BSI) is the recognised UK National Standards Body.

Members: 162 national standards bodies
Standards available: more than 21862
Top 3 standards in popularity:

- ISO 9001 on Quality Management
- ISO 14001 on Environmental Management
- ISO/IEC 27001 on Information Security Management

Usability Standard ISO 9241

- “*Ergonomics of Human System Interaction*”
- Adopts traditional usability categories ...
 - effectiveness
 - can you achieve what you want to?*
 - efficiency
 - can you do it without wasting effort?*
 - satisfaction
 - do you enjoy the process?*
- ISO 9241 was originally entitled “*Ergonomic requirements for office work with visual display terminals (VDTs)*”

Usability Standard ISO 9241

Usability Objective	Effectiveness Measures	Efficiency Measures	Satisfaction Measures
Suitability for the task	Percentage of goals achieved	Time to complete a task	Rating scale for satisfaction
Appropriate for trained users	Number of power features used	Relative efficiency compared with an expert user	Rating scale for satisfaction with power features
Learnability	Percentage of functions learned	Time to learn criterion	Rating scale for ease of learning
Error tolerance	Percentage of errors corrected successfully	Time spent on correcting errors	Rating scale for error handling

Evaluation

Why Evaluate?



“Iterative design, with its repeating cycle of design and testing, is the only validated methodology in existence that will consistently produce successful results.

If you don’t have user-testing as an integral part of your design process you are going to throw buckets of money down the drain.”

Bruce Tognazzini
(founder: Apple Human Interface Group)

<http://www.asktog.com/index.html>

Goals of Evaluation



- Assess the extent of system functionality
- Assess the effect of an interface on a user
- Identify specific problems

Types of Evaluation

- Automated testing
- Testing involving users
 - ‘Controlled settings’
 - usability testing
 - laboratory experiments
 - living labs
 - ‘Natural settings’
 - field studies
(to see how the product is used in the real world)
- Testing not involving users
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - heuristic evaluation
 - review-based evaluation
 - model-based evaluation

Types of Evaluation

Method	Controlled Settings	Natural Settings	Without Users
Observing	✓	✓	
Asking users	✓	✓	
Asking experts		✓	✓
Usability testing	✓		
Modelling			✓

Automated Usability Testing

- Can use software to compare to interface guidelines
- Useful if many interfaces to be examined and little time
- Can use metrics e.g.
 - number of fonts, average size of fonts, average size of buttons, average loading time of graphics
 - might be useful to see how accessible websites are to people with disabilities
 - BUT many aspects need a human check - e.g. appropriate wording? appropriate layout?

Automated Usability Testing

- **Strengths:** a quick method
- **Weaknesses:** many aspects of usability cannot be discovered by automated methods
 - E.g. appropriate wording, labels and layout
 - E.g. can check if webpage has alternative code for graphic (<alt> tag) but not if the text is clear and useful

Types of Evaluation

- Involving users in:
 - ‘Controlled settings’
 - **usability testing**
 - laboratory experiments
 - living labs
 - ‘Natural settings’
 - field studies
(to see how the product is used in the real world)
- Not involving users:
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - heuristic evaluation
 - review-based evaluation
 - model-based evaluation

Usability Testing

- **Aim:**
 - to improve a product's usability
- **Method:**
 - involve real users in the testing
 - give the users real tasks to accomplish
 - enable testers to observe and record the actions of the participants
 - enable testers to analyse the data obtained and make changes accordingly

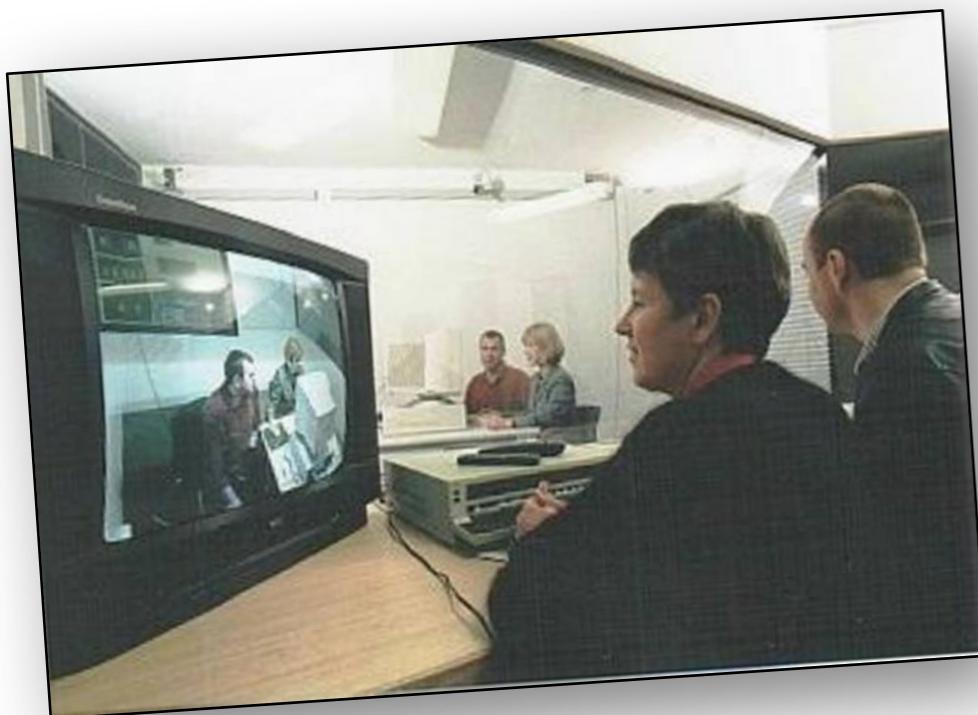
Dumas, J. S., & Redish, J. C. (1993). A Practical Guide to Usability Testing. Norwood, NJ: Ablex.

Types of user-based testing

- **Formative testing**
 - Early testing, might use paper prototype, or WOZ
 - Looking at how user perceives prototype
 - Users may be more prepared to offer feedback at early stage
- **Summative test**
 - When high level design choices made
 - Evaluating effectiveness of choices
- **Validation test**
 - Might compare interface to benchmarks for other interfaces

Usability Testing

- Involves recording performance of typical users doing typical tasks in controlled settings
- Users are observed, timed and recorded
- ...
- data is recorded on video
 - key presses are logged
- The data is used to ...
- calculate performance times
 - identify and explain errors
- User satisfaction can be evaluated using ...
 - questionnaires
 - interviews



Usability Testing



Developer watching videotape of usability test.



Zooming in on the user experience

<http://www.userzoom.com/>
https://www.youtube.com/watch?v=_cneKVGqfNU

What's the customer experience like on

*your
company's*

website or app?



User Testing®

<http://www.usertesting.com>
<https://youtu.be/aqC8NhPPbQw>

Usability Testing

- In ‘usability testing’ the emphasis is on ...
 - selecting representative users
 - developing representative tasks
- Tasks usually last no more than 30 minutes
- The test conditions should be the same for every participant
- An informed consent form ...
 - explains procedures
 - deals with ethical issues



Ethics Applications at TUoS

- Research Integrity and Ethics: Facilitating excellent research
 - <https://www.sheffield.ac.uk/rs/ethicsandintegrity>
- Ethics application process at TUoS
 - <https://www.sheffield.ac.uk/rs/ethicsandintegrity/ethicspolicy/further-guidance/universityprocedure2/uerprocedure>

Usability Testing

Data to collect ...

- time to complete a task
- time to complete a task after a specified time away from the product
- number and type of errors per task
- number of errors per unit of time
- number of navigations to online help or manuals
- number of users making a particular error
- number of users completing task successfully



Usability Testing

- The number of participants is a practical issue and depends on:
 - the schedule for testing
 - the availability of participants
 - the cost of running the tests
- Typically 5-10 participants
(maybe many more for remote web-based/mobile testing)
- Some experts argue that testing should continue until no new insights are gained





- Example: usability testing of Leescircus software
 - Leescircus designed for 6-7 year olds in the Netherlands
 - Example task: Matching pictures that rhyme
 - 70 Dutch children aged 6-7 took part
 - Asked to find problems with software
 - 4 sets of 8 tasks, each child given 1 set
 - Usability evaluators observed children performing tasks
 - Children encouraged to speak comments aloud
 - Time period limited to 30 minutes
 - Findings: need to enlarge clickable objects, clarify the meaning of icons, and improve consistency

Exercise (*work in pairs*)



Identify **three** features
that could be evaluated
by ‘usability testing’

What effectiveness,
efficiency and user
satisfaction measures
would you use?

Types of Evaluation

- Involving users in:
 - ‘Controlled settings’
 - usability testing
 - **laboratory experiments**
 - living labs
 - ‘Natural settings’
 - field studies
(to see how the product is used in the real world)
- Not involving users:
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - heuristic evaluation
 - review-based evaluation
 - model-based evaluation

Laboratory Usability Experiments

- Developers check that the system is usable by the intended user population for their tasks
- ‘**Laboratory usability experiments**’ test hypotheses to discover new knowledge by investigating the relationship between two or more variables
- Method:
 - the independent variable is manipulated by the researcher
 - the dependent variable is what is measured
 - typical experimental designs have one or two independent variables
 - the results are validated statistically and may be shown to be replicable

Laboratory Usability Experiments

- ‘**Different**’ (between group) participants:
 - a single group of participants is allocated randomly to each experimental condition
- ‘**Same**’ (within group) participants:
 - all participants appear in all conditions
- ‘**Matched**’ participants:
 - participants are matched based on expertise, gender, etc.

Types of Evaluation

- Involving users in:
 - ‘Controlled settings’
 - usability testing
 - laboratory experiments
 - living labs
 - ‘Natural settings’
 - field studies
(to see how the product is used in the real world)
- Not involving users:
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - heuristic evaluation
 - review-based evaluation
 - model-based evaluation

Living Labs



“The HomeLab allows Philips to test its new home technology prototypes in the most realistic possible way.”

“Philips researchers carefully watch how tenants are living with these technologies 24 hours a day through tiny cameras, microphones and two-way mirrors that are hidden unobtrusively throughout HomeLab.”

“According to the scientists who developed Philips HomeLab, being able to study people in their natural home environment for long stretches of time will help them to develop better products, faster. It gives them a true sense of how people are interacting with technology beyond the initial “newness” euphoria, and the test subjects act naturally because they are in a comfortable home setting—not a stuffy laboratory.”

SPHERE :: An EPSRC Interdiscipli x

www.irc-sphere.ac.uk

Apps Partnerships & Know... COM 2017/8 Calen... COM - Project Mark... Sample Size Calculator Proposals - Research... Probation - Academ... StudyAbroadRefere... AppCreation Other Bookmarks

Home | About | Get Involved | Research | Blog | Events | News | 100 Homes Study | Partners | People | Contact

sphere
An EPSRC Interdisciplinary Research Collaboration (IRC)

100 HOMES STUDY

ABOUT US

Healthcare is changing - SPHERE is responding. We are developing sensors for the home to diagnose and help manage health and wellbeing conditions. The technology will aid early diagnosis, lifestyle change and the ability of patients to live at home.

[Find out more >](#)
[Introduction film >](#)

HEmiSPHERE Study

LATEST TWEETS

Who wants to be a system architect?
<https://t.co/PvTOZzBpQj>
[read more >](#)

OUR PARTNERS

CONTACT US

NEWS AND EVENTS

Bristolians! Take part in the SPHERE study!

Short films showing the SPHERE sensors installed in homes

An introduction to the SPHERE project by Professor Craddock

SPHERE Animation

© 2015 University of Bristol

<http://www.irc-sphere.ac.uk/>

SPHERE :: An EPSRC Interdisciplinary Research Collaboration (iRC) 100 Homes Study :: SPHERE

www.irc-sphere.ac.uk/100-homes-study

Apps Partnerships & Knowledge COM 2017/8 Calendar COM - Project Management Sample Size Calculator Proposals - Research Probation - Academic StudyAbroadReferees AppCreation Other Bookmarks

Home | About | Get Involved | Research | Blog | Events | News | 100 Homes Study | Partners | People | Contact







SPHERE 100 Homes Study

Background

A vital part SPHERE's research is to test our system of sensors in real homes. To do this we have asked people across Bristol if they'd like to take part in this research. These pages will give you more information about how to take part in this study and what it means to take part in the study.

How to join the study

[Click here](#) for more information on how to join the study.

[How to get involved in the study](#)

If you are interested in having the SPHERE system installed in your home to help with our research, you can find out more [here](#).

[Information for participants](#)

If you are already a participant in the SPHERE study, and would like more information about how the sensors work, what information they collect, or how they will be installed or removed, you can find out more [here](#).

Home Lab



CATCHI
Centre for Assistive Technology and Connected Healthcare



<http://www.cs.vu.nl/en/about-the-department/intertain-lab/index.asp>

<https://www.youtube.com/watch?v=nirrSw2SrEM#t=27>

Types of Evaluation

- Involving users in:
 - ‘Controlled settings’
 - usability testing
 - laboratory experiments
 - living labs
 - ‘Natural settings’
 - **field studies**
(to see how the product is used in the real world)
- Not involving users:
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - heuristic evaluation
 - review-based evaluation
 - model-based evaluation

Field Studies

- ‘**Field studies**’ are conducted in natural settings (“*in the wild*”)
- The aim is to understand ...
 - what users do naturally
 - how technology impacts them
- Field studies are used to ...
 - identify opportunities for new technology
 - determine design requirements
 - decide how best to introduce new technology
 - evaluate technology in use

Field Studies

- Field study data is collected using ...
 - observation (*audio/video logging*)
 - interviews
- Data is usually presented as ...
 - anecdotes
 - excerpts
 - critical incidents
 - patterns
 - narratives
- The aim is to show how the products are being appropriated and integrated into their surroundings

Types of Evaluation

- Involving users in:
 - ‘Controlled settings’
 - usability testing
 - laboratory experiments
 - living labs
 - ‘Natural settings’
 - field studies
(to see how the product is used in the real world)
- Not involving users:
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - heuristic evaluation
 - review-based evaluation
 - model-based evaluation

Cognitive Walkthrough

- Proposed by Polson *et al.*
 - evaluates design on how well it supports user in learning task
 - usually performed by expert in cognitive psychology
 - expert ‘walks through’ design to identify potential problems using psychological principles
 - forms used to guide analysis
- For each task walkthrough considers ...
 - what impact will interaction have on user?
 - what cognitive processes are required?
 - what learning problems may occur?
- Analysis focuses on goals and knowledge: does the design lead the user to generate the correct goals?

Cognitive Walkthrough

- In order to perform a ‘cognitive walkthrough’, you need ...
 - a specification or prototype of the system
 - a description of the task the user is to perform on the system
 - a complete written list of the actions needed to complete the task
 - an indication of levels of experience and knowledge of the users
- For each action, the evaluators must answer the following questions ...
 - *“Is the effect of the action the same as the user’s goal at that point?”*
 - *“Will users see that the action is available?”*
 - *“Once users have found the correct action, will they know it is the one they need?”*
 - *“After the action is taken, will users understand the feedback they get?”*

Types of Evaluation

- Involving users in:
 - ‘Controlled settings’
 - usability testing
 - laboratory experiments
 - living labs
 - ‘Natural settings’
 - field studies
(to see how the product is used in the real world)
- Not involving users:
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - **heuristic evaluation**
 - review-based evaluation
 - model-based evaluation

Heuristic Evaluation



Jakob Nielsen: The first pioneering “King of Usability”
<http://www.nngroup.com/people/jakob-nielsen/>

- Proposed by Nielsen in 1994
 - ‘usability criteria’ (*heuristics*) are identified
 - design examined by 3-5 evaluators to see if these are violated
- Example heuristics
 - “*system behaviour is predictable*”
 - “*system behaviour is consistent*”
 - “*feedback is provided*”
- Heuristic evaluation ...
 - can find ~75% of overall usability problems
 - is flexible
 - is cheap

Nielsen, J. (1994). Heuristic evaluation, Usability Inspection Methods. New York: John Wiley.

Heuristic Evaluation



- Nielsen's ten heuristics (broad rules of thumb)
- Does the system ...
 - always keep users informed about what's going on?
 - use language that is familiar to the user?
 - support 'undo' and 'redo'? (emergency exit)
 - follow platform standards and conventions?
 - make it difficult to make errors?
 - make objects, actions and options visible? (minimise memory load)
 - allow users to tailor frequent actions? (accelerators for experienced users)
 - minimise irrelevant or rarely needed information?
 - express error messages in plain language? (suggest a solution)
 - provide help and documentation?

Heuristic Evaluation



- Shneiderman's 8 golden rules of interface design
 - Strive for consistency
 - Cater to universal usability
 - Offer informative feedback
 - Design dialogs to yield closure
 - Prevent errors
 - Permit easy reversal of actions
 - Support internal locus of control
 - Reduce short-term memory load

Heuristic Evaluation

Usability problems are rated by evaluators on a scale of 0-4 ...

0 = no problem at all

1 = cosmetic problem only

2 = minor usability problem

3 = major usability problem

4 = catastrophic problem



Types of Evaluation

- Involving users in:
 - ‘Controlled settings’
 - usability testing
 - laboratory experiments
 - living labs
 - ‘Natural settings’
 - field studies
(to see how the product is used in the real world)
- Not involving users:
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - heuristic evaluation
 - **review-based evaluation**
 - model-based evaluation

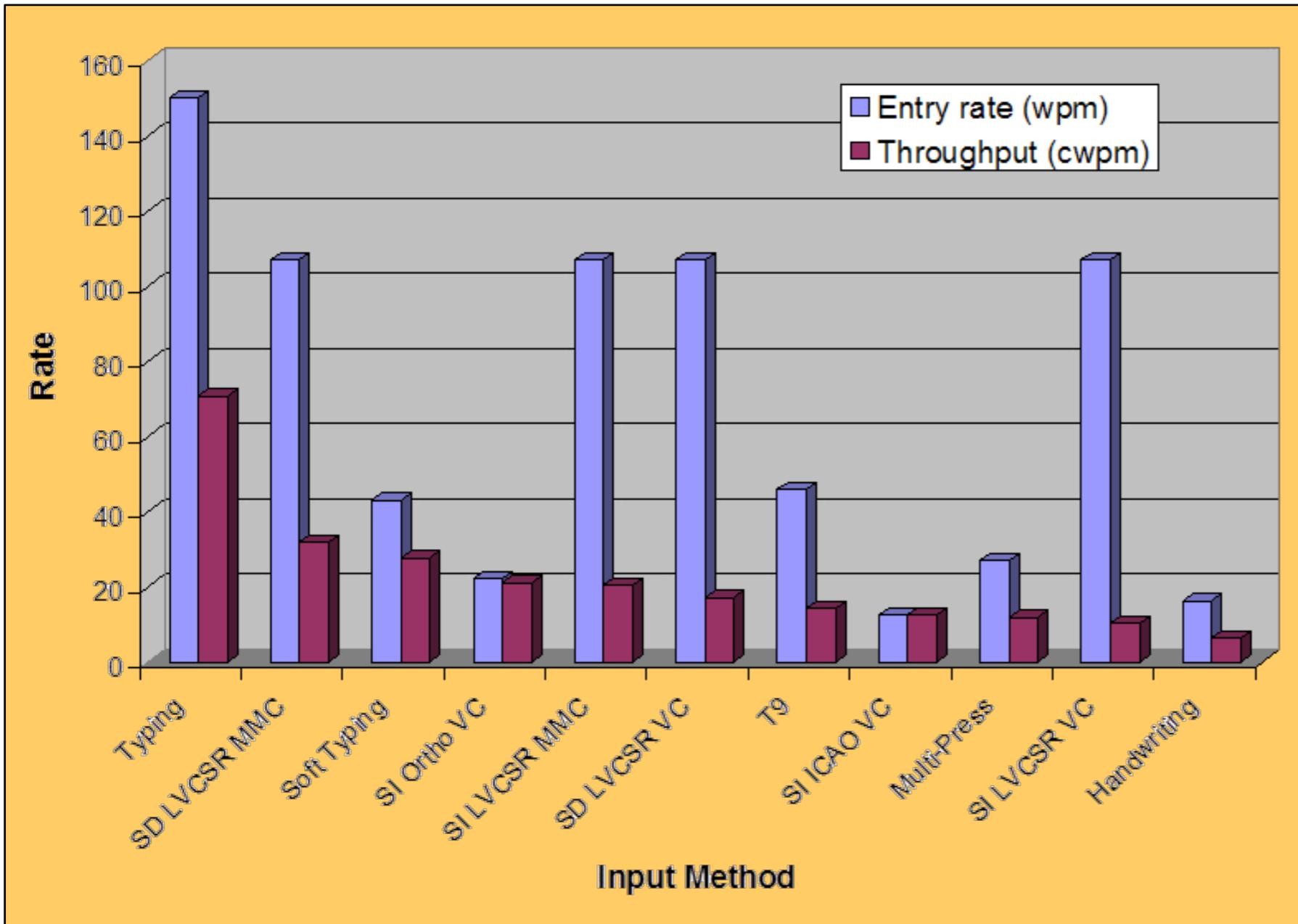
Review-Based Evaluation

- Results from the literature are used to support or refute parts of a design
- Cheaper than (re)running experiments
- Care is needed to ensure results are transferable to new design
- Reviewers must ensure that there is a match between ...
 - the experimental designs chosen
 - the user populations used
 - the analyses performed
 - the assumptions made

Types of Evaluation

- Involving users in:
 - ‘Controlled settings’
 - usability testing
 - laboratory experiments
 - living labs
 - ‘Natural settings’
 - field studies
(to see how the product is used in the real world)
- Not involving users:
 - ‘Experts’ critique, predict, analyse and model aspects of the interface
 - cognitive walkthrough
 - heuristic evaluation
 - review-based evaluation
 - **model-based evaluation**

Model-Based Evaluation



Using Fitts' Law
to predict data
entry rates on a
PDA

Choosing an Evaluation Method

- What stage in the life-cycle? *design vs. implementation*
- What style of evaluation? *laboratory vs. field*
- How objective? *subjective vs. objective*
- What types of measure? *qualitative vs. quantitative*
- What level of information? *high level vs. low level*
- What level of interference? *obtrusive vs. unobtrusive*
- What resources are available?
 - time
 - subjects
 - equipment
 - expertise

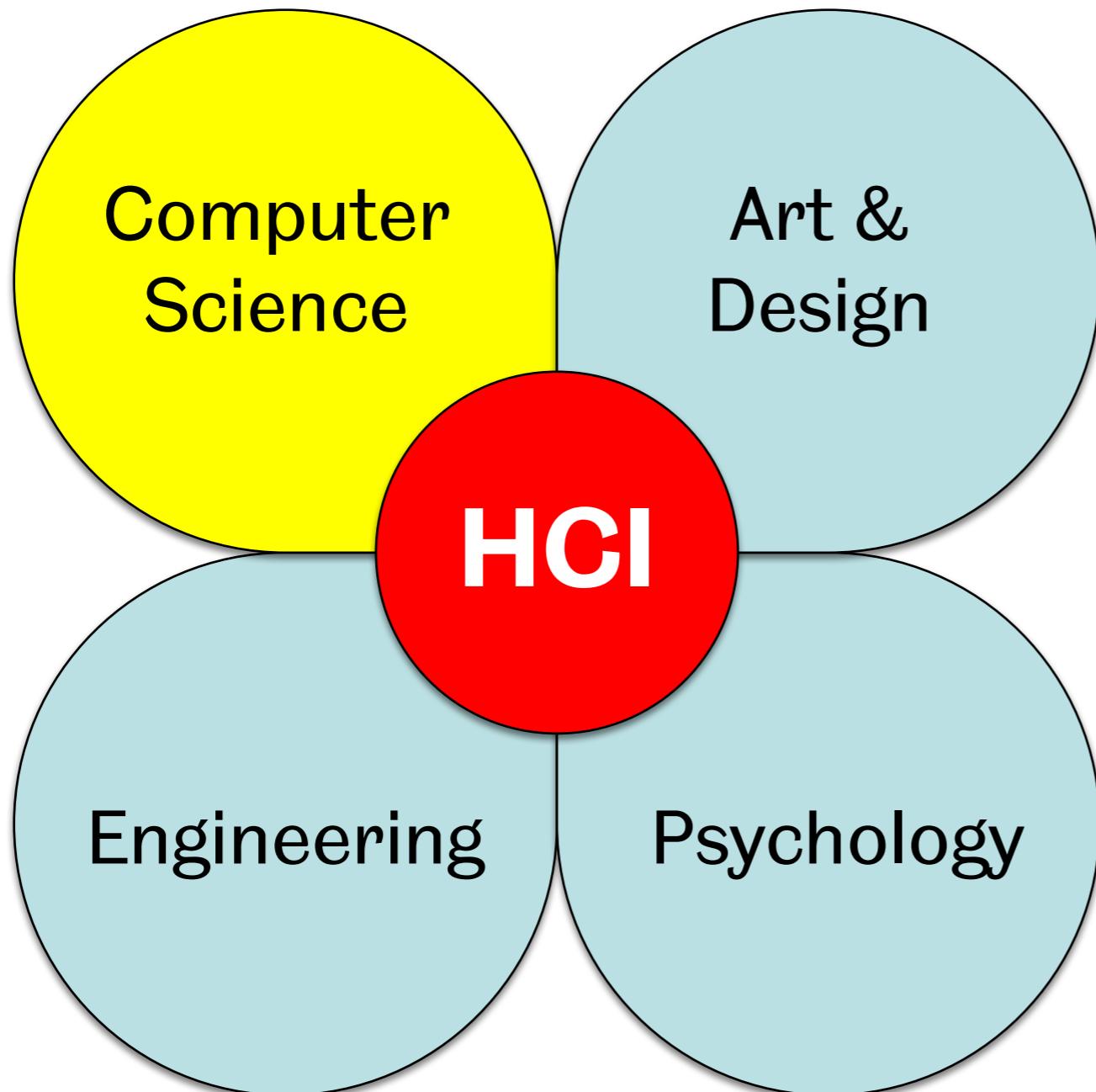
Exercise (*work in pairs*)

- You have developed a prototype of a '**theatre booking system**' aimed at reducing queues at the box office
- Choose an appropriate evaluation method and identify ...
 - the participants
 - the techniques used
 - representative tasks to be examined
 - measurements that would be appropriate
 - an outline plan for carrying out the evaluation
- Be prepared to share your solution with the class

Summary

- Good design is about ...
 - usability
 - ensuring that systems are *accessible* to all
 - ensuring that designs are *acceptable* for the people and contexts in which they will be used
- Designers need to **evaluate** their designs with people and involve people in the design process
- Access to interactive systems for all people is an important right
- Usability is concerned with balancing the PACT elements in a domain
- Acceptability is concerned with ensuring designs are appropriate to contexts of use

User Interface Design



**multi-
disciplinary!!!**

Feedback ?



Thank You