

# Understanding and serving users

Usability - how to define it and  
how to measure it

# Usability- a key concept

- Defining usability
  - Semantic, Feature and Operational
- ‘Usability engineering’
  - criteria
  - contextual variables
  - task analysis

# What is usability?

- Simple semantic definitions:
  - ‘user-friendliness’?
  - ‘ease-of-use’?
  - ‘ease-of-learning’?
  - ‘transparency’
  - ‘intuitive to use’

# Semantic definitions

- Usable means 'easy to use'...

So what does 'ease of use' mean?

Semantic definitions tend to circularity for design purposes

# What is usability (part 2)?

- Can we define usability by features?
  - Windows, icons, menu, pointer (WIMP)?
  - Same look&feel as popular resources?
  - Is designed according to user requests?
  - Meets Nielsen's principles of design?
  - Has been tested on real users?

# Attribution Fallacy

- The attribution fallacy suggests usability is a quality of an interaction that is determined by the presence or absence of specific features present in the interface.
- This attribution leads to an over-reliance on guidelines and prescriptive rules for design

# Can we say **NOTHING** about features?

- We can generalize within contexts
- There are many guidelines which work for **MOST** users **MOST** of the time
  - e.g., image polarity and reading speed
  - e.g., choice is better than forced path
  - e.g., quick better than delayed feedback
- But we cannot define usability reliably on these

# So what is usability?

- Need a definition that is:
  - explanatory
    - defines the construct meaningfully
  - unambiguous
    - is not open to multiple interpretations
  - has utility for information resource design
    - can inform designers and drive user-centered design process

# Operational definition

Usability refers to an application's "capability (in human functional terms) to be used easily and effectively by the specified range of users, given specified training and support, to fulfill the specified range of tasks within the specified range of environmental scenarios" Shackel, (1991)

# Operational definition

Usability (of an application or resource) refers to the effectiveness, efficiency, and satisfaction with which specified users can achieve specified goals in particular environments

ISO Ergonomics requirements, ISO 9241 part 11: Guidance on usability specification and measures.

# Effectiveness

The extent to which users can achieve their task goals.

Effectiveness measures the degree of accuracy and/or completion

e.g., if desired task goal is to locate information on a web site then:

Effectiveness = success of user in locating the correct data

# Effectiveness can be a scale or an absolute value

- If the outcome is ALL or NOTHING then effectiveness is an absolute value
  - User either locates info or does not...
- If outcome can be graded, (user can be partially right) then effectiveness should be measured as a scale
  - As a %, or a score from 1 (poor) to 5 (complete)
- Scale should be determined by evaluator in conjunction with developers and users

# Quality?

- Some tasks do not have a definitive correct answer:
  - creative production (writing, design)
  - information retrieval
  - data analysis
  - Research
  - management
- Effectiveness alone misses something...

# Efficiency

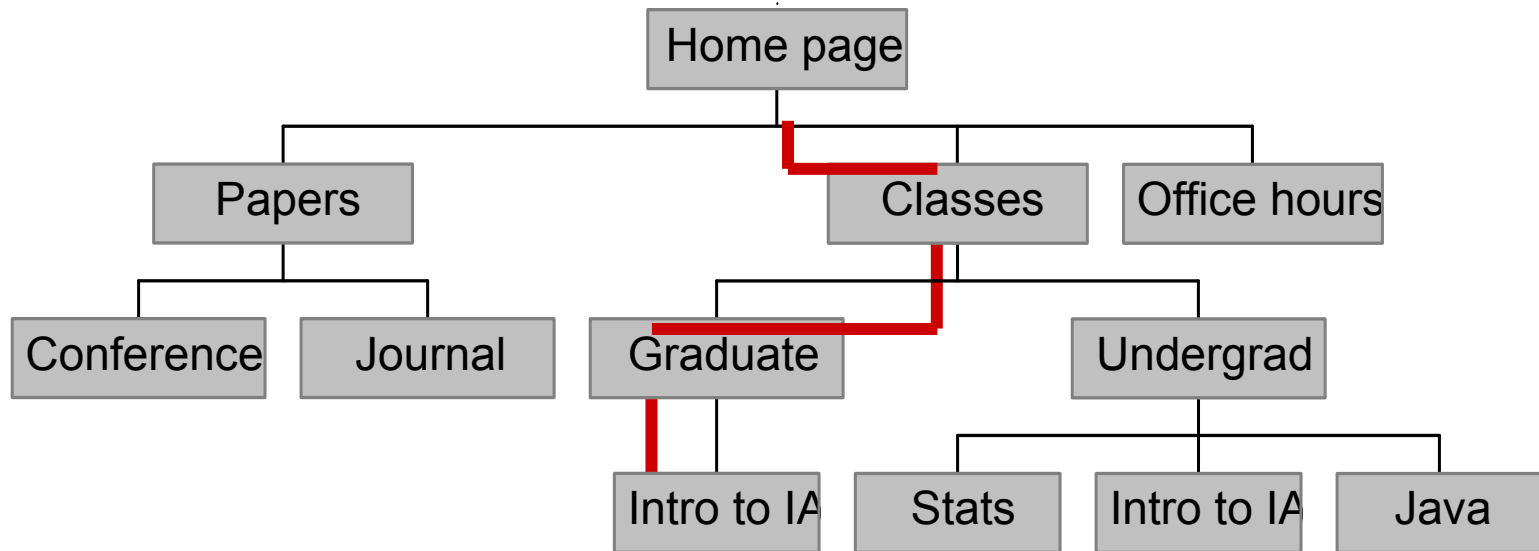
- Measures resources used to perform task
  - i.e., time, effort, cost,
- In case of Web site use, efficiency might equal time taken to complete a task or the navigation path followed etc.

# Efficiency of using a redesigned web site

- Time taken to complete task
  - Compared across tasks, across users or against a benchmark score
- Number of steps taken
- Number of deviations from ideal path

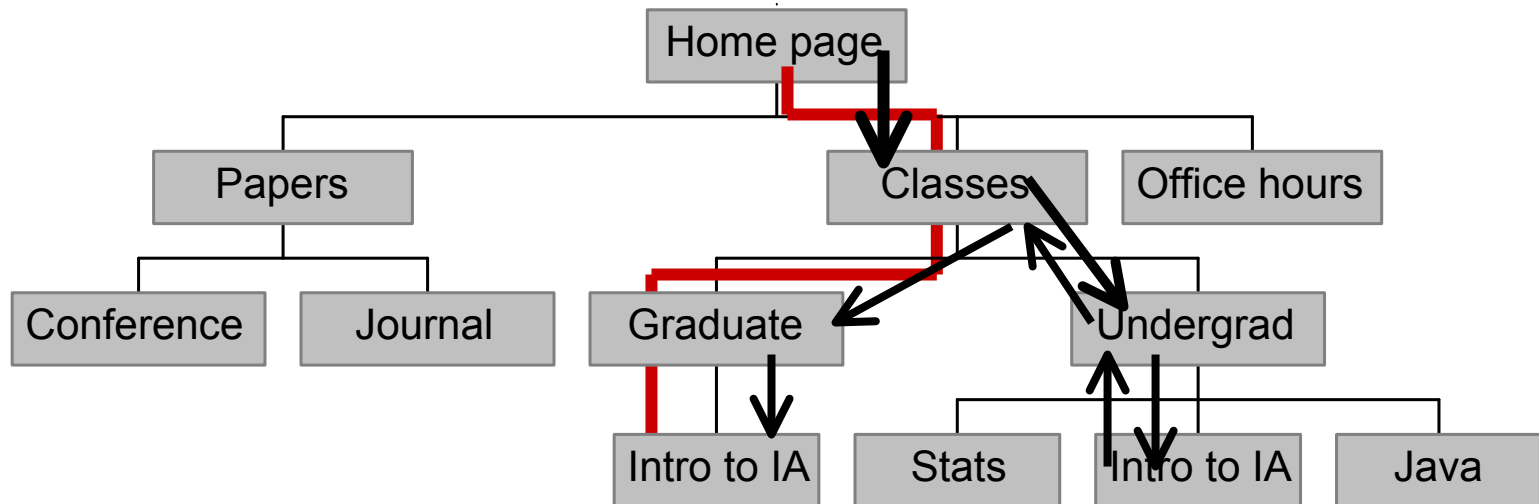
Such variables are frequently highly positively correlated - but they needn't be.

# Efficiency in path analysis



Ideal path: 3 steps Understanding and Serving Users 8

# Efficiency in path analysis



Actual to ideal user navigation: 7:3 steps

# But is it efficiency that users want?

- The push to efficiency is symptomatic of an engineering-oriented approach
  - Who determines efficiency?
  - Are path deviations always inefficient?
  - Is time equally weighted by user, designer or owner?
- Suggests a need for negotiation beyond typical usability tests

# Satisfaction

- Measures the affective reaction (likes, dislikes, attitudinal response) of users to the application/resource
- Assumed to be influenced but not the same as effectiveness or efficiency e.g.,
  - 2 resources with equal effectiveness, and efficiency, may not be equally satisfying to use
  - or What users like might not be what they need!

# Basis for satisfaction?

- Positively influenced by effectiveness and efficiency
- Also
  - Personal experience with other technologies?
  - Working style?
  - Manner of introduction?
  - Personality of user?
  - Aesthetics of product?

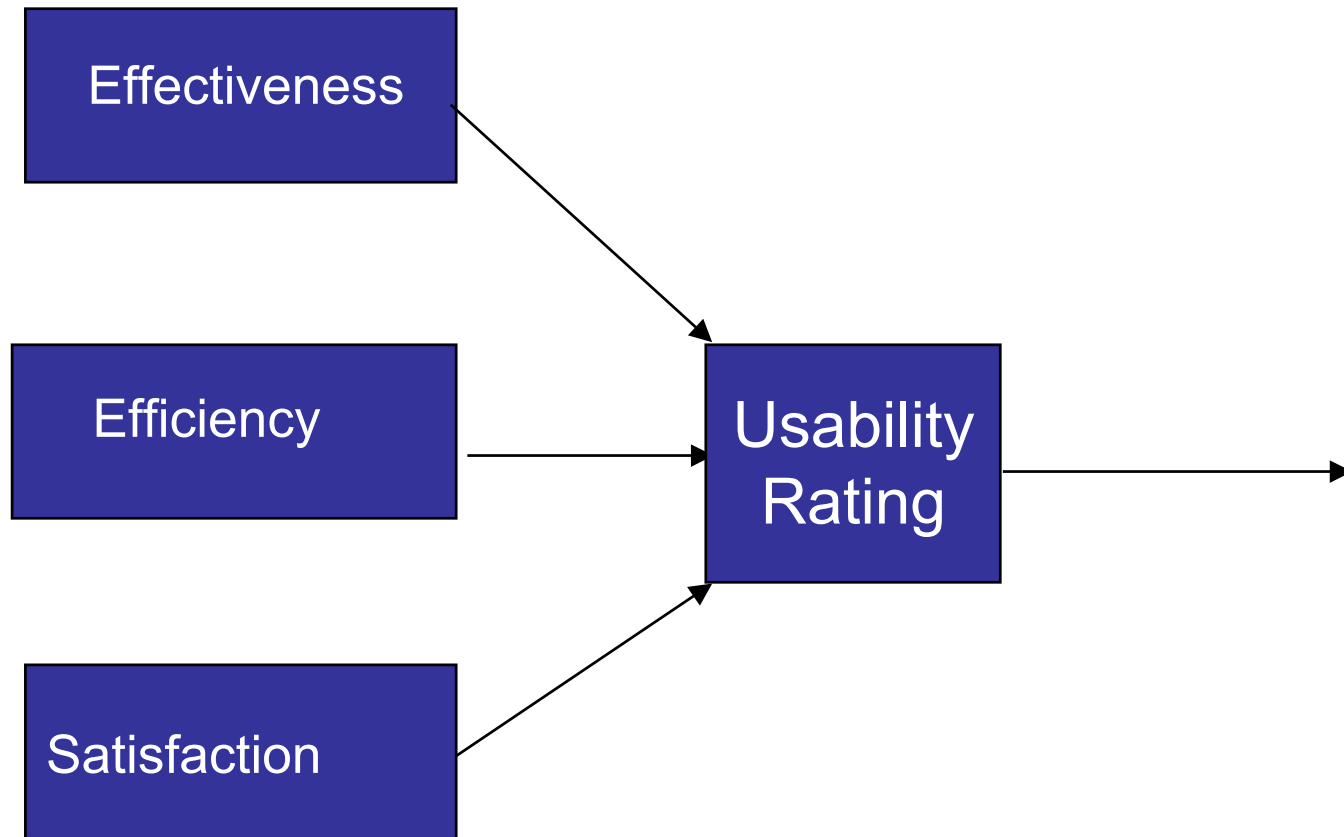
# Satisfaction is important

- Good usability studies recognize this

## But satisfaction is not enough....

- People often like what they don't use well
- What about empowerment, challenge etc?

# Determinants of usability rating



# Recap

- Usability is a 'negotiable' construct
- Should be expressed in measurable terms
- Measures should reflect design questions
- Capture performance and affect
  - effectiveness
  - efficiency
  - satisfaction
- Compare across contexts with caution!

# Setting usability criteria

“Resource X is usable to the extent that 70% of users, with no additional training, can perform all tasks with 95% accuracy, 25% faster than existing version, and report at least equivalent satisfaction”

# Or.....

“Resource X is usable to the extent that 80% of users, with 2 days training, can perform 90% of routine tasks with >90% accuracy, as efficiently as with the existing version, and report increases in satisfaction”

# Instead of....

“Information resource X is usable”  
(meaningless statement for our purposes)

“This new service is more usable than the  
previous version”

(begs the question...”More usable in what  
sense? And for whom? And where?”

# Or

- Resource Y is usable to the extent that 3/4 of users, after 5 mins of training, can perform all tasks with 100% accuracy, in 2 minutes (or less) per task, and report mean satisfaction levels of 4 on a 5 pt scale.

# Who sets the usability criteria?

- Owners/Purchasers
  - can be basis for contract
- Designers
  - basis for design targets
- Evaluators
  - provide context/limits of generalization for evaluation
- Users
  - Key stakeholders with privileged knowledge

# How are criteria derived?

- User analysis
- Task analysis
- Contextual analysis

# User analysis

- Determine key variables:
  - work and task skills
  - Task and resource experience
  - training
  - support
  - working practices
  - and many other variables (see previous lectures)

# Task analysis

The process of observing, describing, and decomposing tasks into their constituent components and analyzing the human, technological and environmental resources required for the completion of each component.

# Task analysis questions

- What does the user see?
- What decisions does a user make?
- What must a user know?
- How does a user get help?
- How does a user recover from errors?
- What physical acts must be accomplished?

# Task analysis methods

- Contextual inquiry
- Observation
- Interview
- Verbal protocol analysis
- Self trial
- Formal descriptive method (e.g. GOMS analysis)

# Situation analysis

- Outlines the physical and social context of use:
  - Location
    - Home, office, shop, car, street etc.
  - Relationship to other users
    - Collaboration, recipient, passive/active
  - Socio-technical environment

# Output

- These analyses provide guidance to determining the required levels of:
  - Effectiveness
  - Efficiency
  - Satisfaction
- And the training, tasks, and environment in which the application or resource will be used

# Why definition matters?

- What you define usability to be, determines what you design into being
- Understanding how usability is determined improves the quality of its evaluation and the interpretation of test results
  - Definition drives evaluation

# Two reasons why evaluation matters:

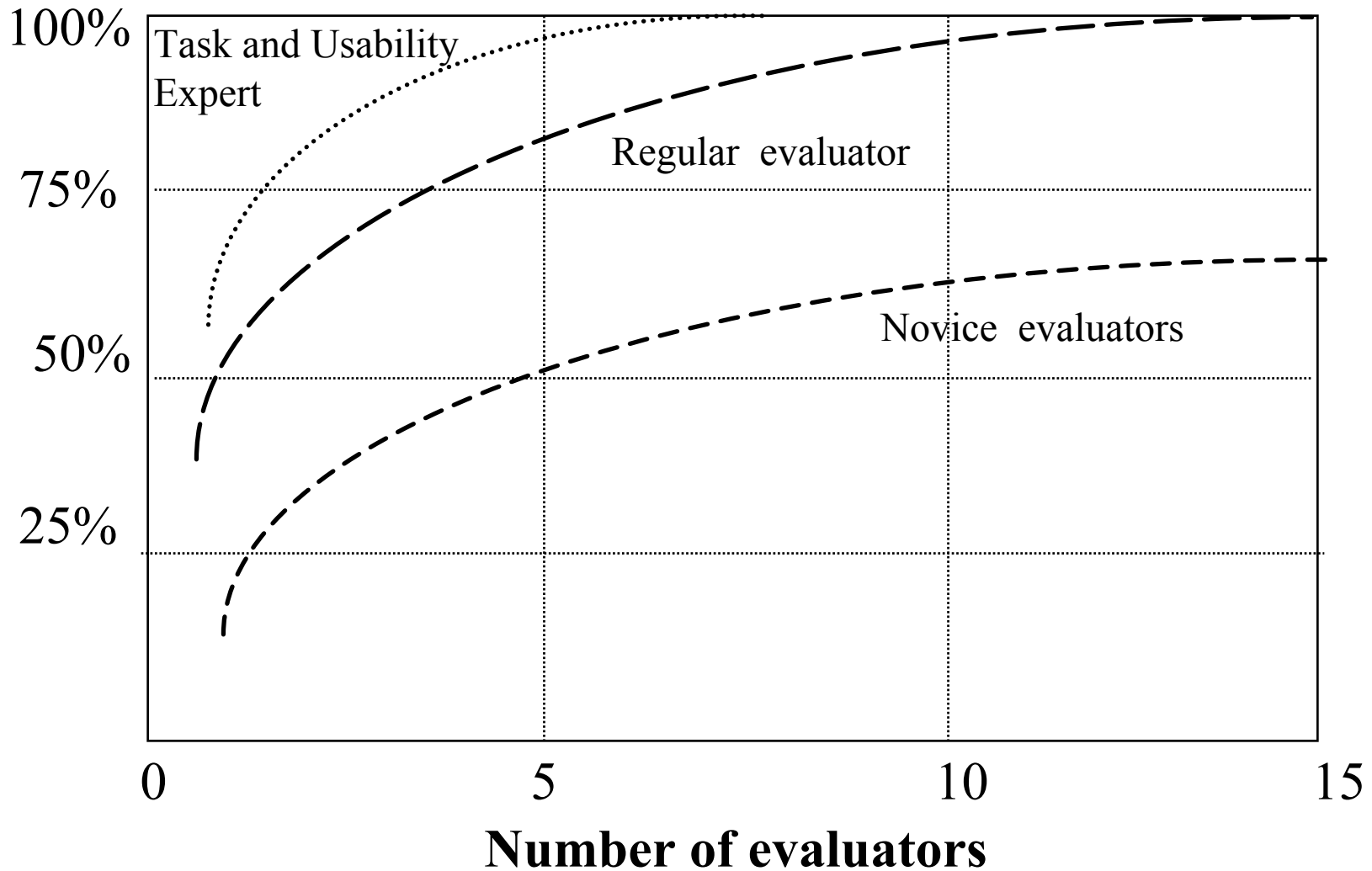
- Designers make poor predictions of use
- Users make poor predictions of use

# But all evaluations are not the same.....

- Evaluators vary in their ability to identify problems
- Evaluation methods are not equally useful

# Comparing evaluations - Nielsen 1993

Problems Found



# Framework for Usability Evaluation

- Approach and Type
  - Approach refers to source of data
    - User, Expert, or Model
  - Type refers to purpose of evaluation
    - Diagnostic (Formative) or Metrication (Summative)
- Any evaluation method is a combination of *approach* and *type*

# Evaluation Approach

The approach defines the *source* of the data i.e., where does the evaluator gain the data about usability?

- from real users? (**User-based**)
- from usability experts or self evaluation? (**Expert-based**)
- from the application of a formal theory or model? (**Model-based**)

# User-based approach

- Involves one or more users
- Who complete one or more tasks
- In an appropriate environment

In formal usability tests, the user, task, and environmental characteristics (*contextual variables*) must match those in which the product will be used.

Many informal variants exist

# Advantages of User-based approach

- Most valid data (true measure of usability)
- Most reliable data (measure is stable)
- Provides the most convincing data
- Can gain responses to low-fidelity prototype

# Disadvantages of User-based approaches

- Can be expensive
  - Time
  - Resources
- Requires patience
- Must be planned (to be useful)
  - in terms of tasks, likely scenarios etc.
- Does not logically determine re-design

# Expert-based Approaches

- Involves the evaluator (usability expert) using the system
- Use can be structured or casual
- Expert notes potential problems
- May employ pre-determined criteria
- Expert tries to see the interface from the point of view of the user

# Typical Expert-based Scenario (1)

- Heuristic evaluation (Nielsen 1993)
  - Take Nielsen's guidelines and check interface against them
    - Simple and natural dialog
    - Provide feedback
    - Provide good error messages.....
  - Make 2 passes through, once for a screen by screen check of principles, once more for inter-screen flow

# Typical Expert-based Scenario (2)

- Cognitive Walkthrough
  - “a hand simulation of the cognitive activities of the user” Polson (1992)
  - Take task analysis and ‘walkthrough’ all activities that a user must perform, checking for consistency, comprehensibility etc.

# Advantages of Expert-based Evaluations

- Can be quick
- Cheaper than user testing
- Can lead to diagnosis
- Can be employed at all stages of design
- No special facilities required

# Disadvantages of Expert-based Evaluations

- Who determines 'expertise'
- Expertise in task is also important (Nielsen 1993 study)
- Experts disagree!
- Predicting user behavior is difficult

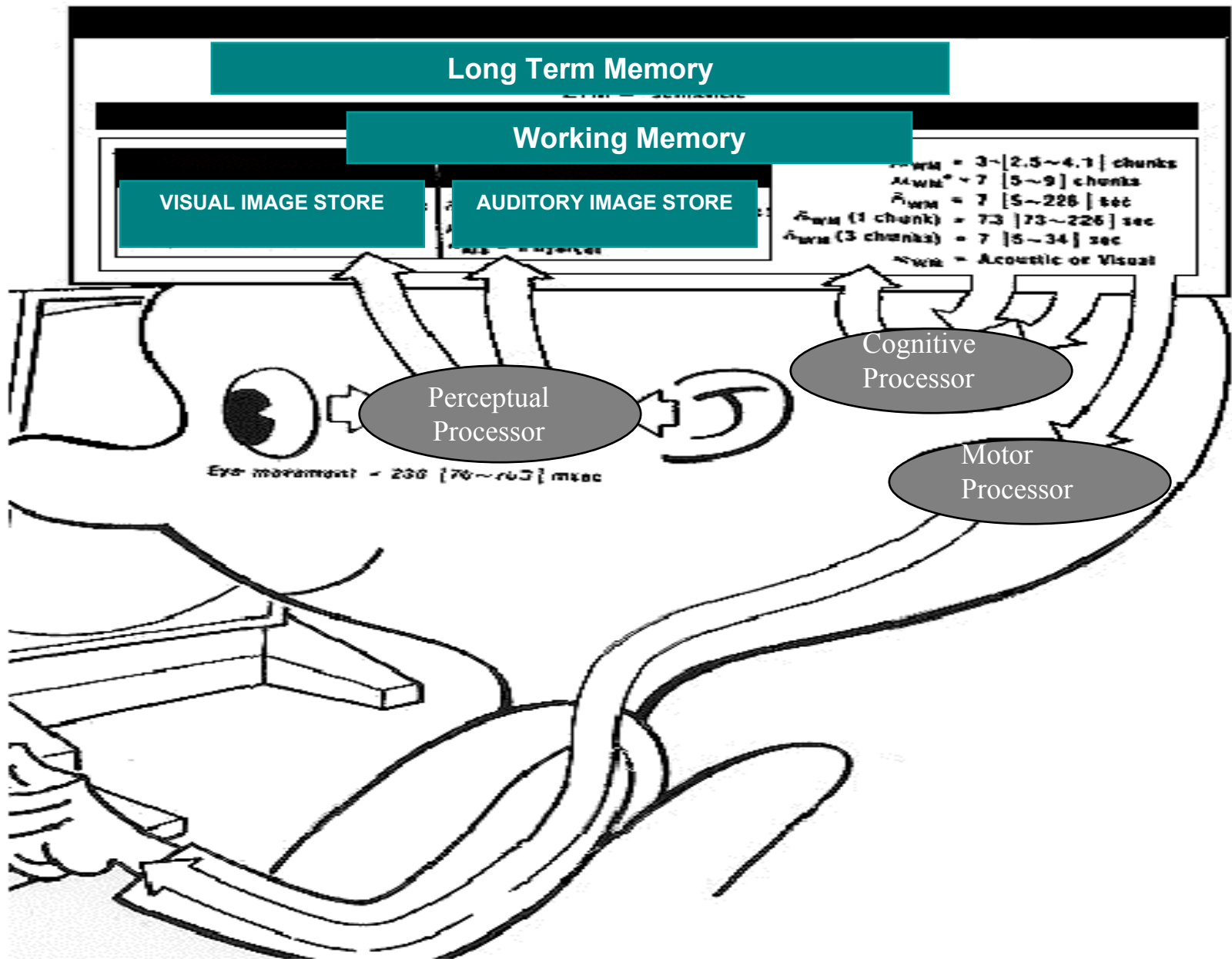
# Model-based Evaluations

- Involve the calculation of the match between outputs of task analysis and the user interface design
- Often based around formal languages
- Rarely used in HCI

# Typical Model-based Scenario

- Evaluator measures the complexity of system e.g., in terms of the production rules a user must learn
- Evaluator predicts time to completion for expert performer using task-analysis based methods (e.g., GOMS, Card et al 1983)

(see Preece et al (1995) Human-Computer Interaction)



Long Term Memory

Working Memory

VISUAL IMAGE STORE

AUDITORY IMAGE STORE

Perceptual Processor

Cognitive Processor

Motor Processor

Eye movement = 238 [70~400] msec

Capacity = 3~[2.5~4.1] chunks  
Duration = 7 [5~9] chunks  
Duration = 7 [5~226] sec  
Duration (1 chunk) = 73 [73~226] sec  
Duration (3 chunks) = 7 [5~34] sec  
Duration = Acoustic or Visual

# Cognitive parameters

From: Olson and Olson, 1990

Name	Action	Time (msec)
K	Enter a keystroke	230
M	Point with the mouse	1500
Hm	Move handsto mouse	360
P	Perceive	100
R	Retrieve from memoy	1200
Ex	Execute a mental step	70
Ch	Choose among methods	1250

# Save file <MYFILE> in MS Word

Initial homing of hand	(Hm=0.36 sec)
Move cursor to file menu, + Retrieve from memory	(M =1.5) (R=1.2)
Click file menu	(K =0.23)
move down menu + retrieve from memory	(M=1.5) (R=1.2)
click 'save as'	(K=0.23)
System reacts	(Sys=1.2),
user types name and hits return	(R =1.2 + (nk= 6(0.23)) <u>(K=0.23)</u> )
total:	10.23 secs

# Advantages of Model-based Evaluations

- Can be performed at the specification stage
- Cheap
- Enables maximum scope for change

# Disadvantages of Model-based Evaluations

- Complex to apply without some training
- Too few evaluators trained in these methods
- Limited range of predictive power
  - GOMS predicts only expert, error-free performance completion times for non-discretionary tasks

# Type of Usability Evaluation

- Diagnostic/Formative
  - Evaluation is aimed at uncovering shortcomings in design
  - May use all approaches described already
  - This is most common
  - Are all usability evaluations diagnostic (formative)?

# Type of Usability Evaluation (2)

- Metric/certification/summative
  - Seeks to certify or measure product against pre-defined criteria(standards, certification)
  - Usually relies on user-based tests
  - More common for non-usability testing
  - Even this type of evaluation can be used for diagnostic purposes in on-going design iterations

# Things to try.....

- Consider the usability issue in terms of your own experiences
- Derive a guideline for user interface design that applies in all contexts of use?
  - do any exist?
- Think about what usability would mean for some of your routine technologies
  - ATM, VCR, Email etc.
- Readings....readings....readings....