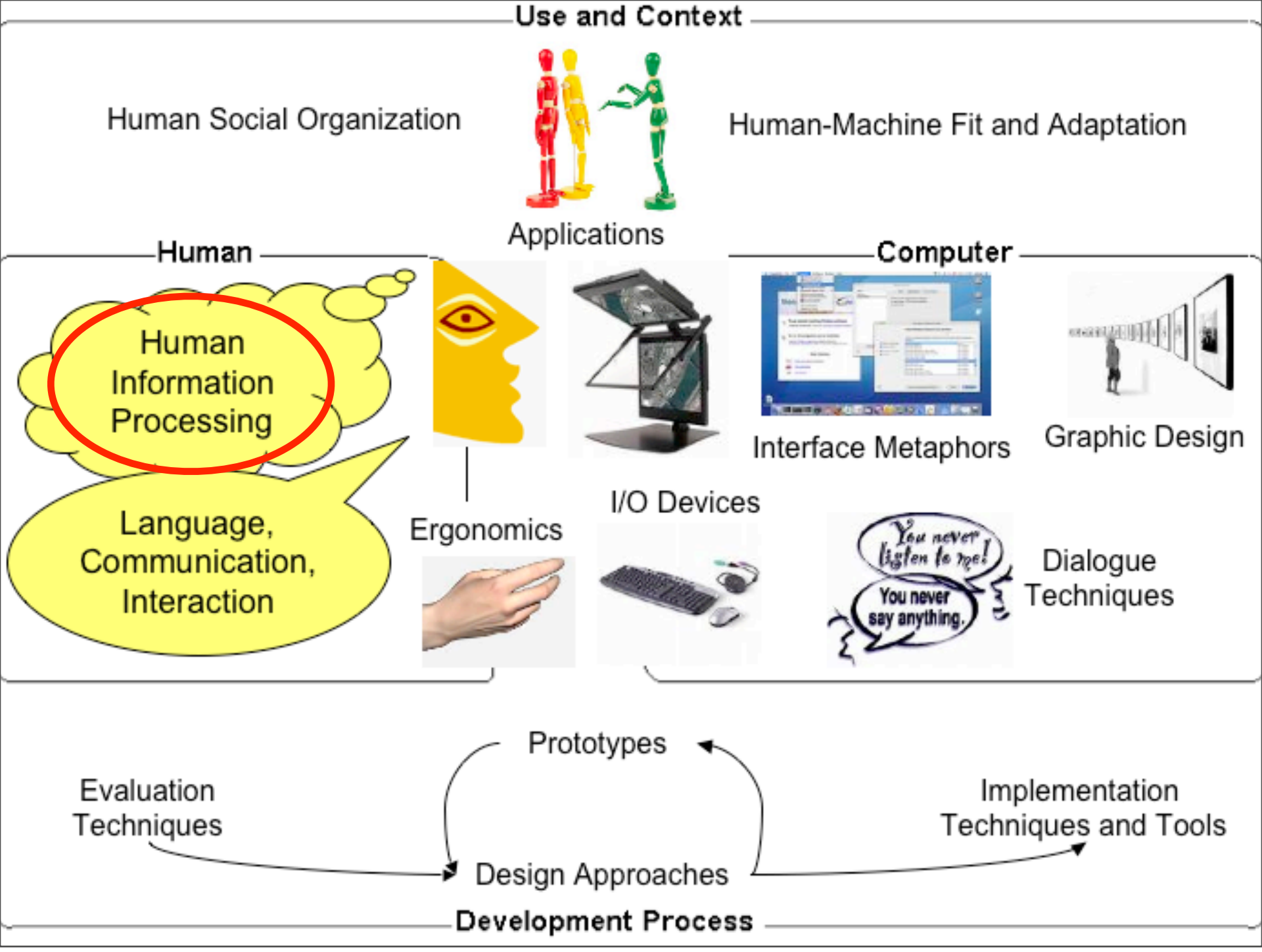
Human Cognitive Models

The Human (Chapter 1)


Use and Context
Human Social Organization
Human-Machine Fit and Adaptation
Applications
Human
Computer
Human Information Processing
Language, Communication, Interaction
Ergonomics
I/O Devices
Interface Metaphors
Graphic Design
You never listen to me!
You never say anything.
Dialogue Techniques
Prototypes
Evaluation Techniques
Implementation Techniques and Tools
Design Approaches
Development Process

Introduction

- Why is the Nintendo Wii so much fun to use?
- Because it minimizes the distance between the human's capabilities and what we want the machine to do!

Overview

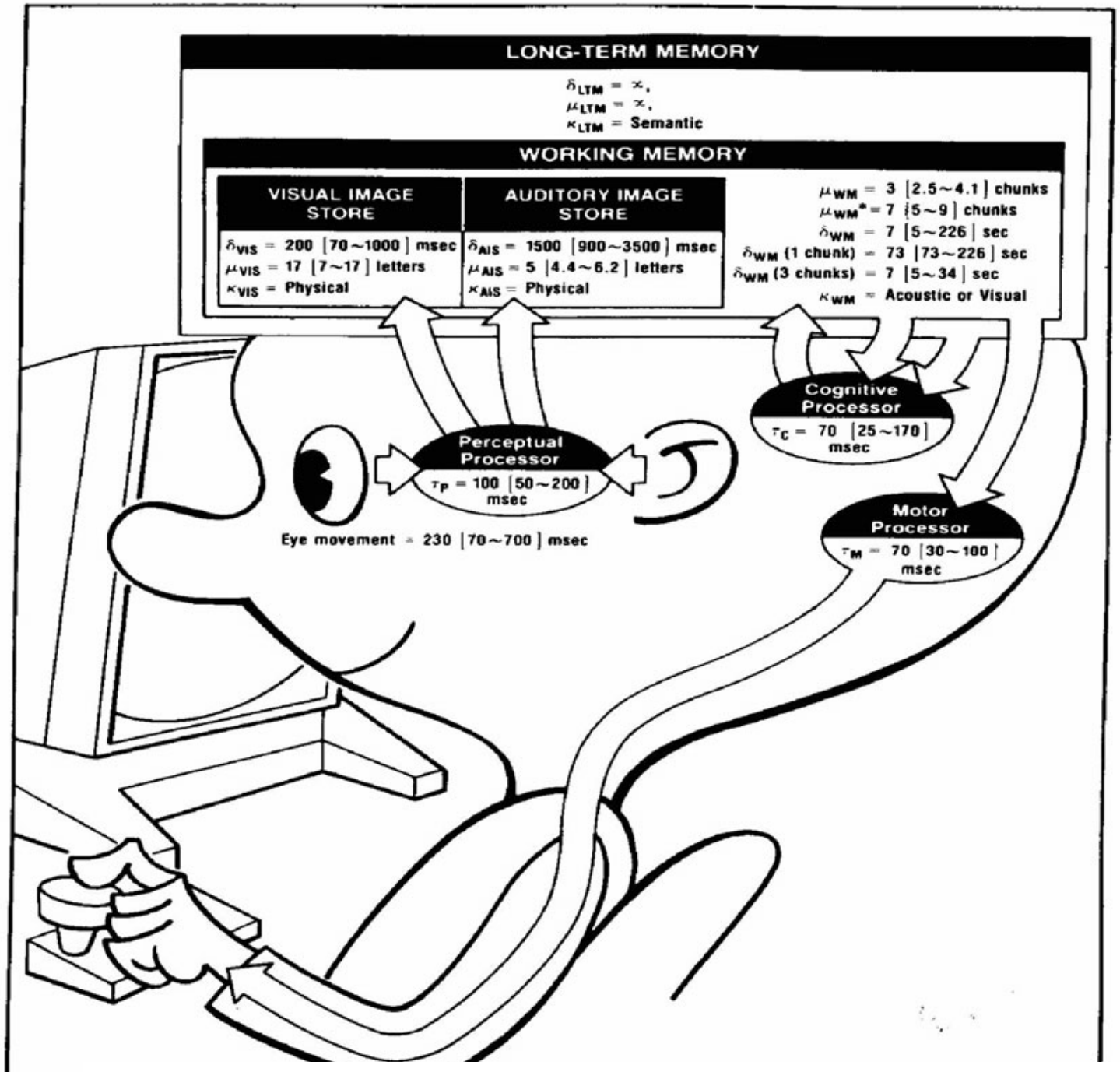
- What you'll learn in this lecture:
  - How do humans work?
  - What are the limitations of humans?
  - What should we watch for when we design?

Why are we learning about this?

- Humans are only human, after all.
- Limited capacity for processing information.
- Must be taken into account if good interfaces are to be designed.

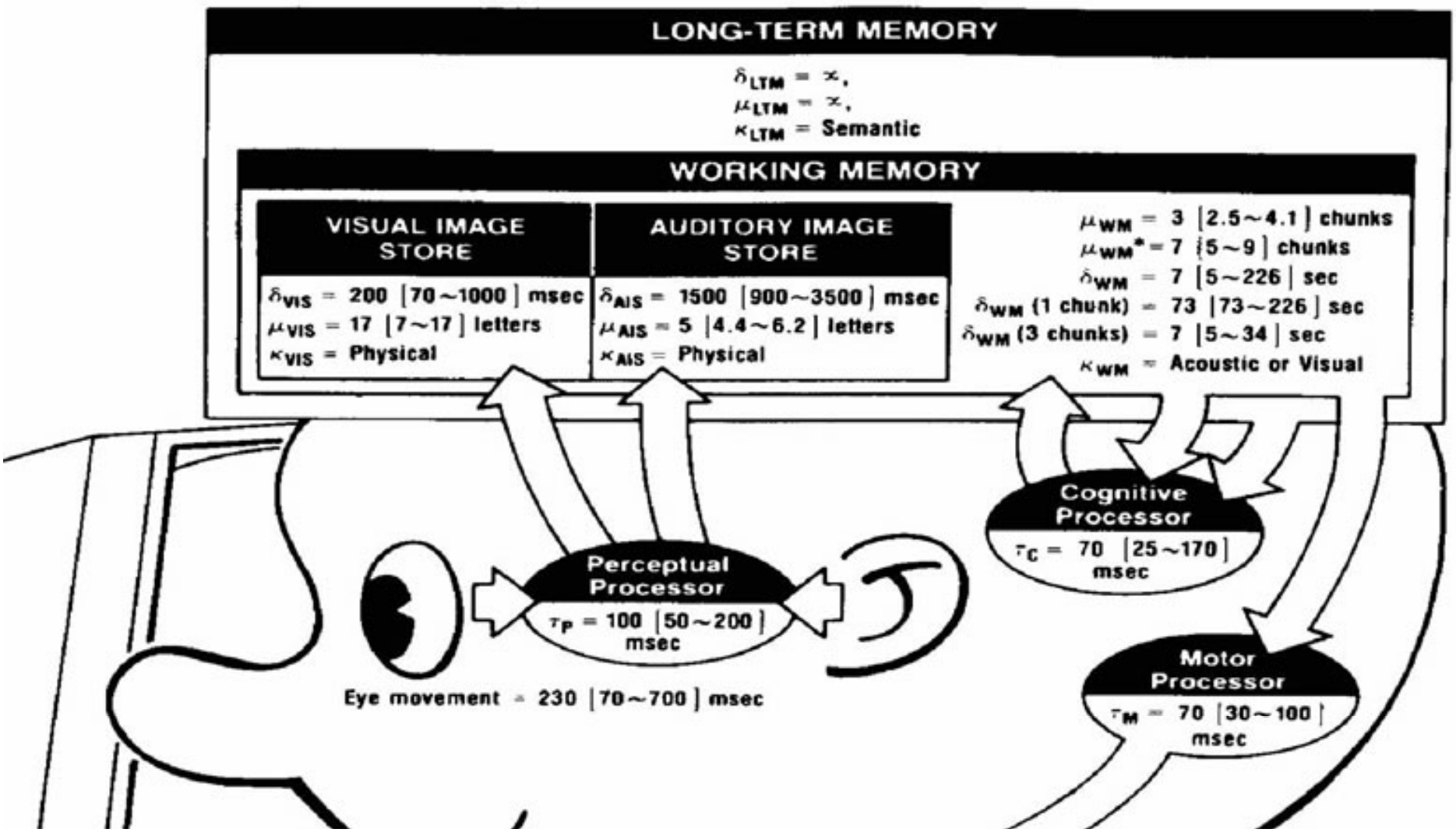
Human Behavior (Timescale)

10^7 (months) 10^6 (weeks) 10^5 (days)	Social	Social Behavior
10^4 (hours) 10^3 10^2 (minutes)	Rational	Adaptive Behavior
10^1 10^0 (seconds) 10^{-1}	Cognitive	Immediate Behavior
10^{-2} 10^{-3} (milliseconds) 10^{-4}	Biological	

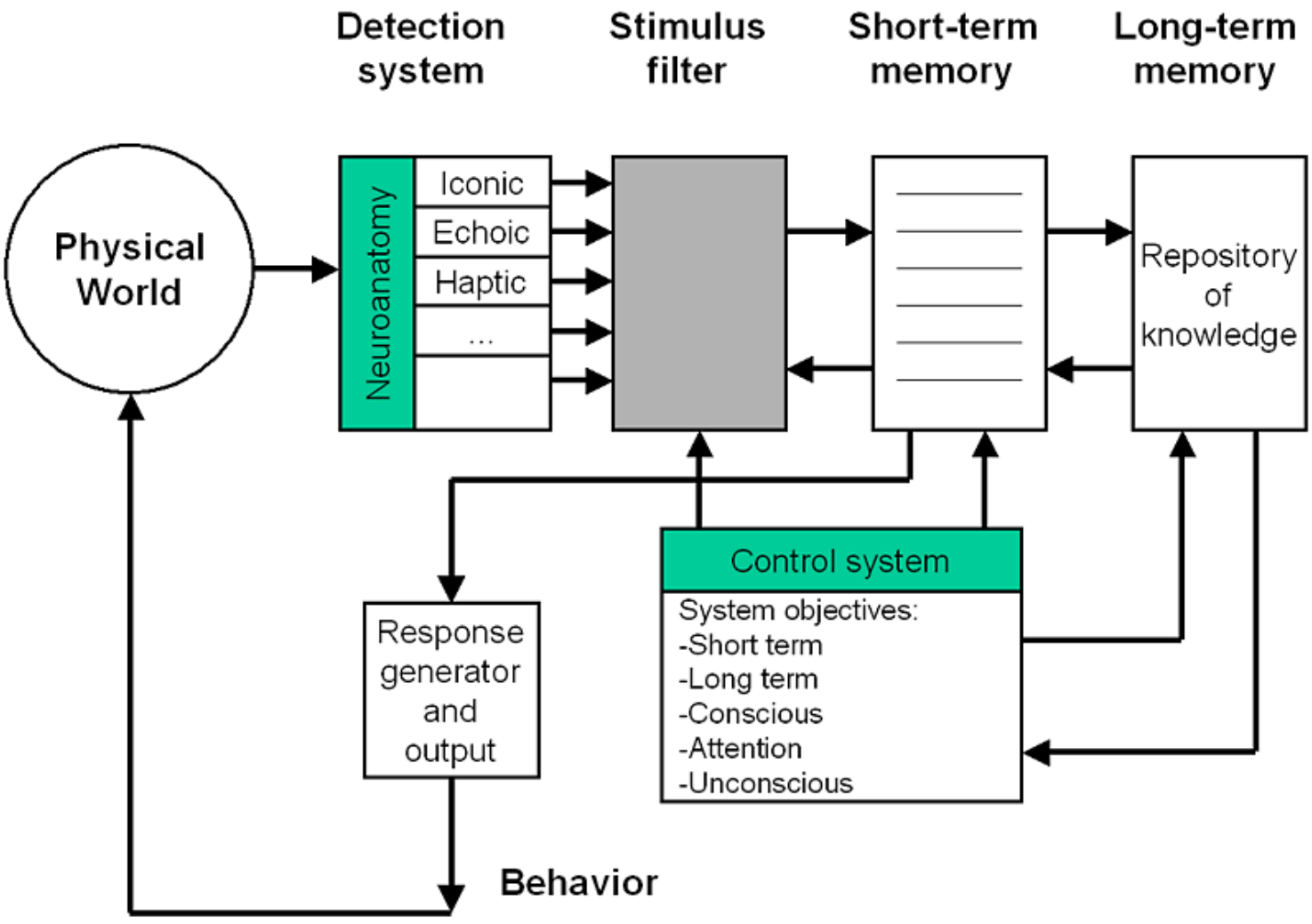


Card, Moran, & Newell, The Psychology of Human-Computer Interaction (1983)

The Human Information Processor



The Human Information Processor



Model of the human brain inspired by the digital computer (R. L. Solso, Cognitive Psychology)

Human Information Processing

- Vision
- Hearing
- Touch
- Movement
- Memory
- Thinking: Emotion

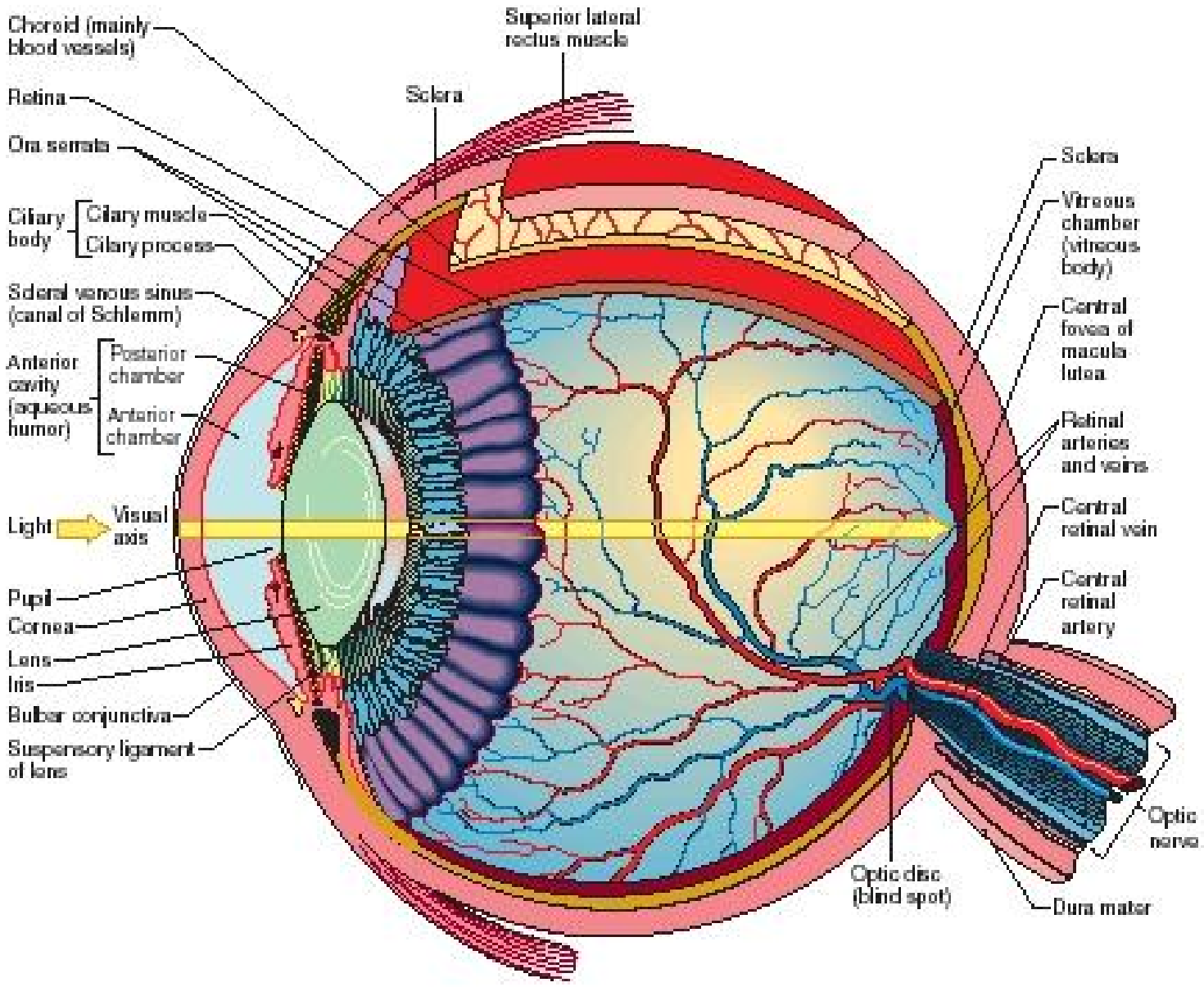
The Human

- Information input/output
  - Visual, auditory, movement
- Information stored in memory
  - Sensory, short-term, long term
- Information processed and applied
  - Reasoning, problem solving, skill, error
- Emotion influences human capabilities
- And each person is different!

Iconic Input (Vision)

- Two Stages:
  - Physical reception of stimuli
    - Receive light and transform it into electrical energy
    - Light reflects from objects; their images are focused upside-down on retina
    - Retina contains rods for low-light vision and cones for color vision
    - Ganglion cells detect pattern and movement

Vertical section of the right eye, shown from the nasal side



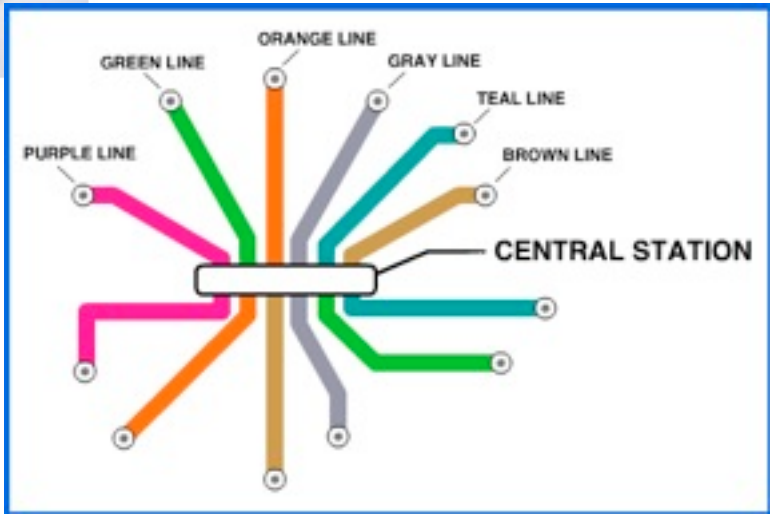
Iconic Input: Brightness

- Brightness
  - subjective reaction to levels of light
  - affected by luminance of object
  - measured by just noticeable difference
  - visual acuity increases with luminance as does flicker

Iconic Input: Color

- Color
  - Hue (property of the wavelength of light, i.e. color)
  - Intensity (lighter/darker)
  - Saturation (purity of the hue in a color)
    - The higher the saturation, the purer the color.
  - Cones sensitive to color wavelengths
  - Blue acuity is lowest
  - 8% males and 1% females colour blind

Color-blindness

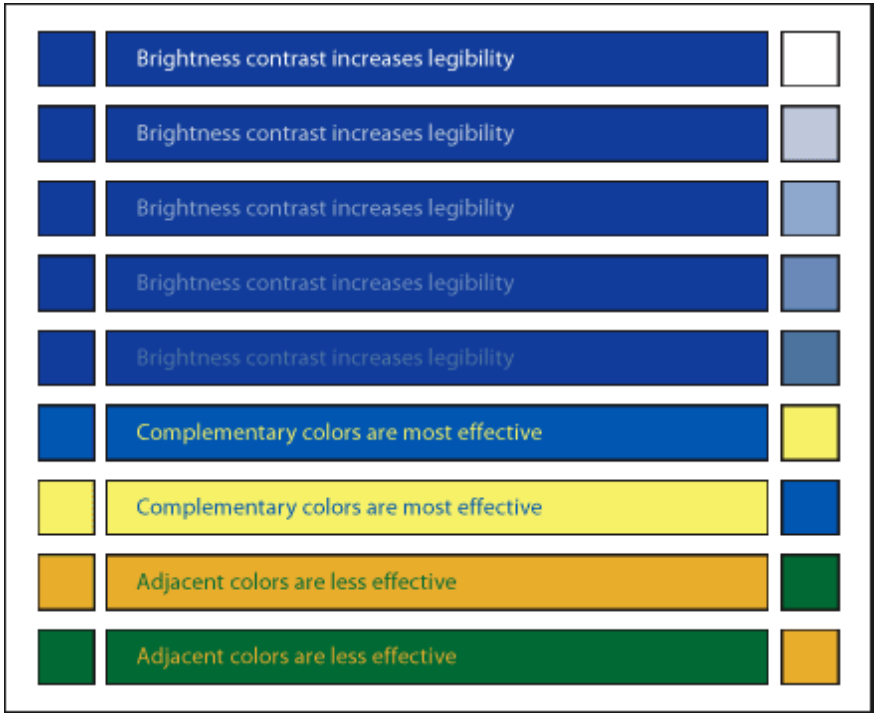


Color-blindness

COMP homepage, as it appears to normal people, and people with red-green color-blindness.
(From http://colorfilter.wickline.org/)

Color for Text

- Color differentiation is affected by both brightness and hue.
- When possible, go for maximum contrast in brightness (black and white) and hue (yellow and purple)
- Negative contrast -- dark characters on a light screen
  - Provides higher luminance than positive contrast
  - Improves reading from computer screen



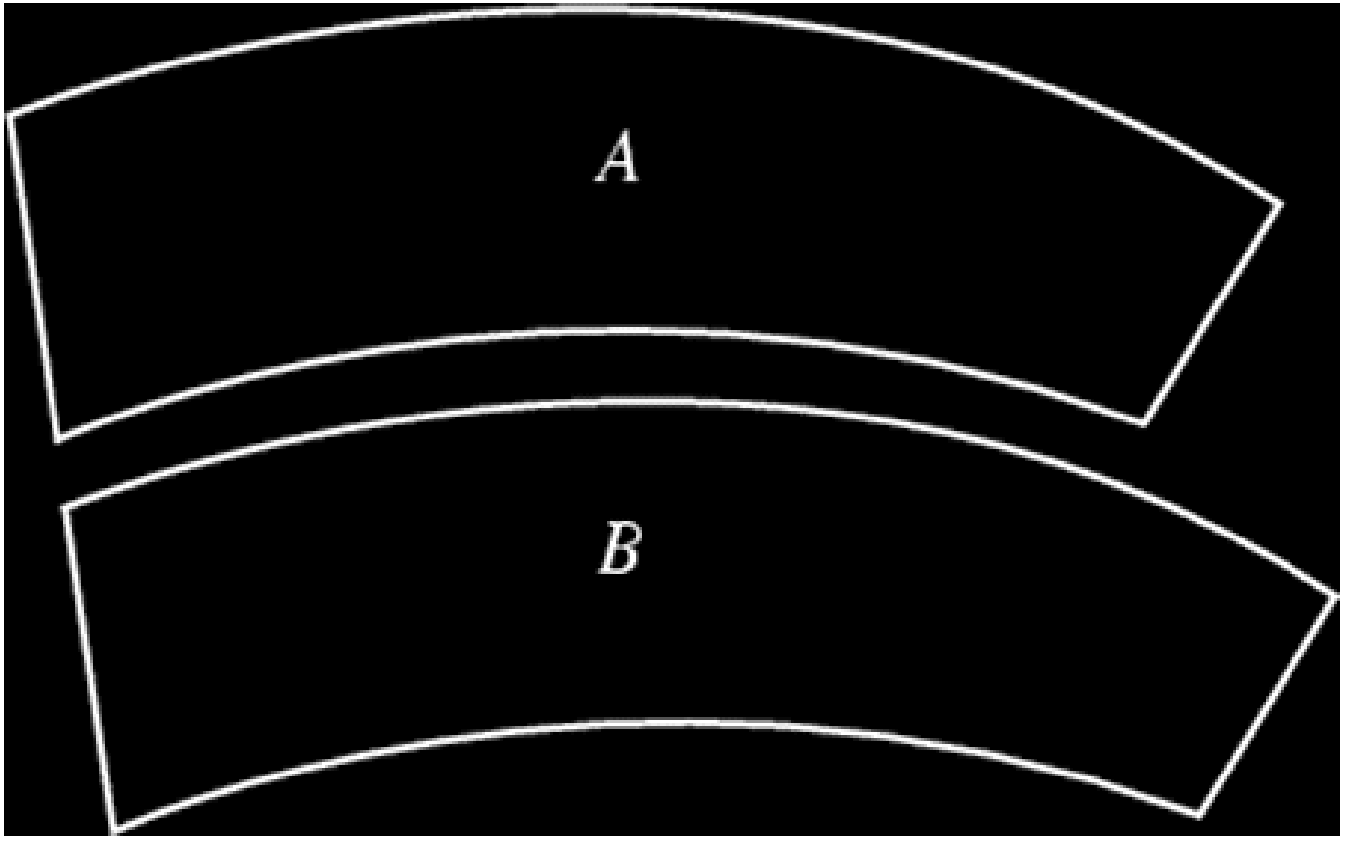
Iconic Input: Interpreting the Signal

- Size and depth
  - Visual angle indicates how much of view is occupied by the object
  - Related to size and distance from eye
- Visual acuity is ability to perceive fine detail
- Law of size constancy means that familiar objects are perceived as constant size in spite of changes in visual angle
- Cues like overlapping help perception of size and depth.

Interpreting the Signal: Perception

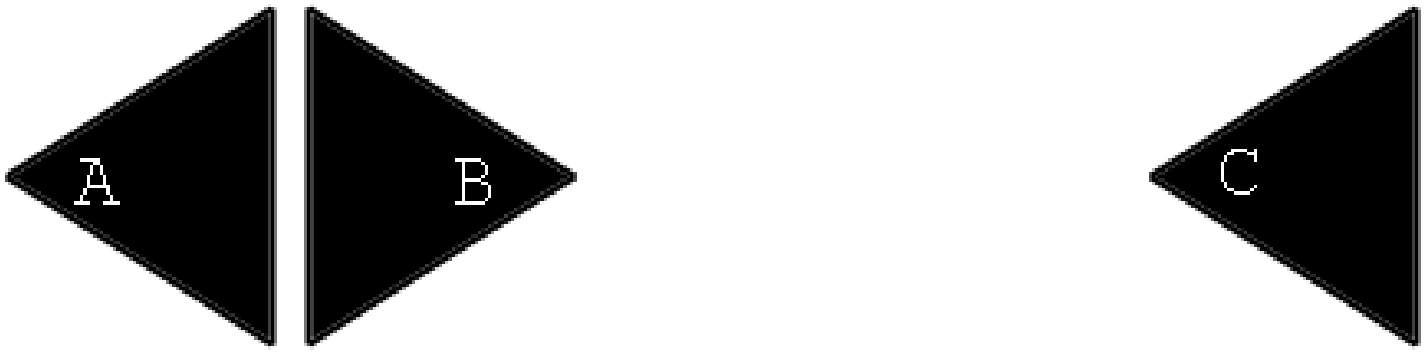
- The human brain interprets the signal coming from the eyes and perceives them as forms and patterns.
  - Compensates for movement and changes in luminance.
  - Context is used to resolve ambiguity.
- Optical illusions sometimes occur due to over-compensation.

Which one is bigger?

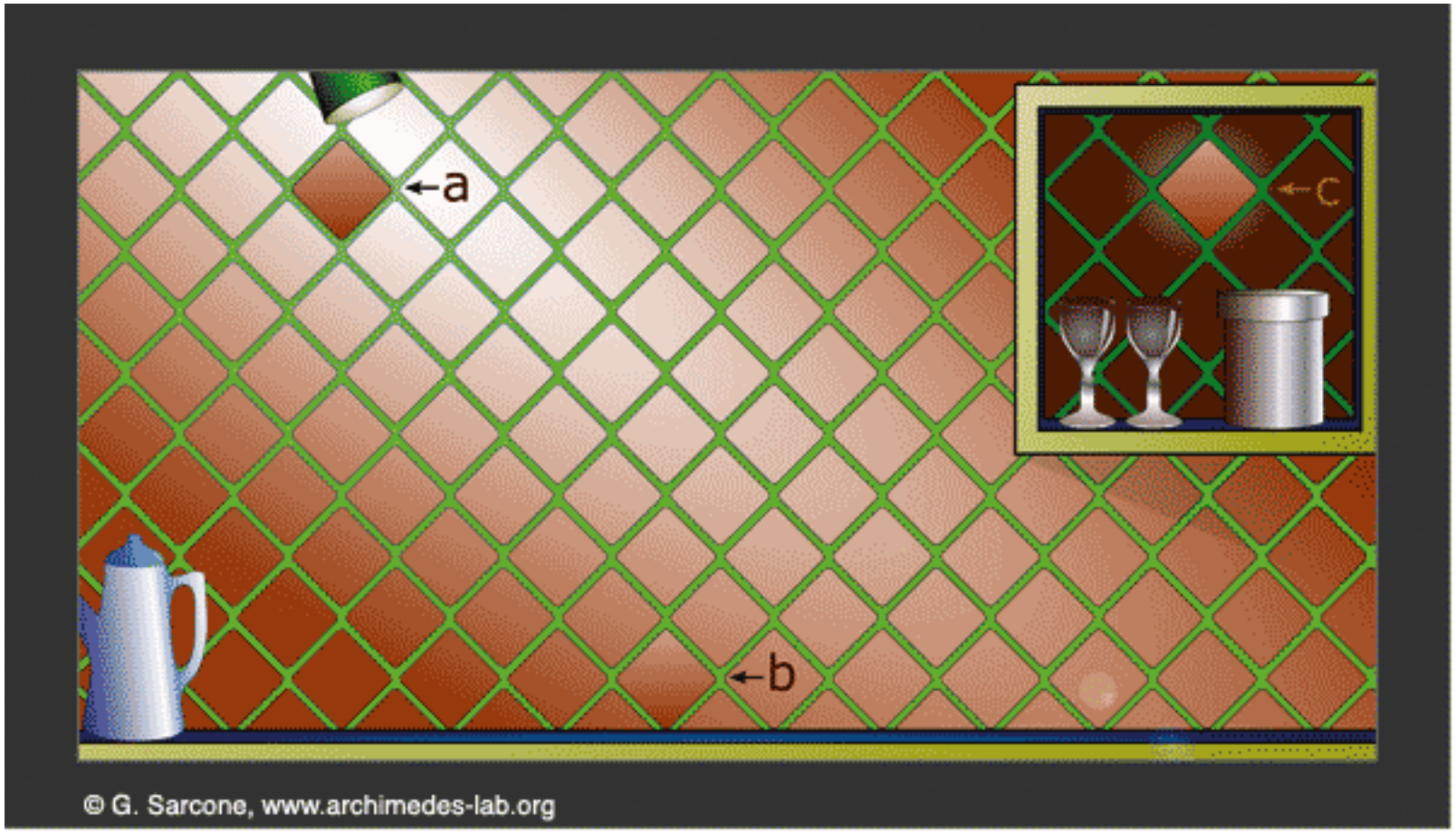


Jastrow Illusion

AB <> BC?



Which square is the darkest?



Interpreting the Signal: Perception

- Gestalt Perception
- Fundamental Principle: Law of Pragnanz (consciousness)
  - We tend to order our experience in a manner that is regular, orderly, symmetric, and simple.
- Laws hypothetically allow us to predict the interpretation of sensation

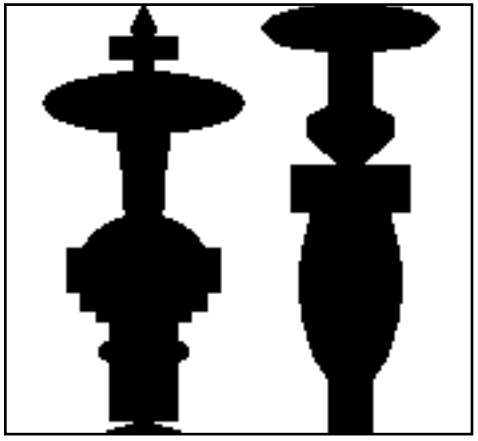
The Law of Similarity

- Items that are similar tend to be grouped together.

- Most people will see columns of squares and circles (rather than rows of alternating shapes)

Law of Symmetry

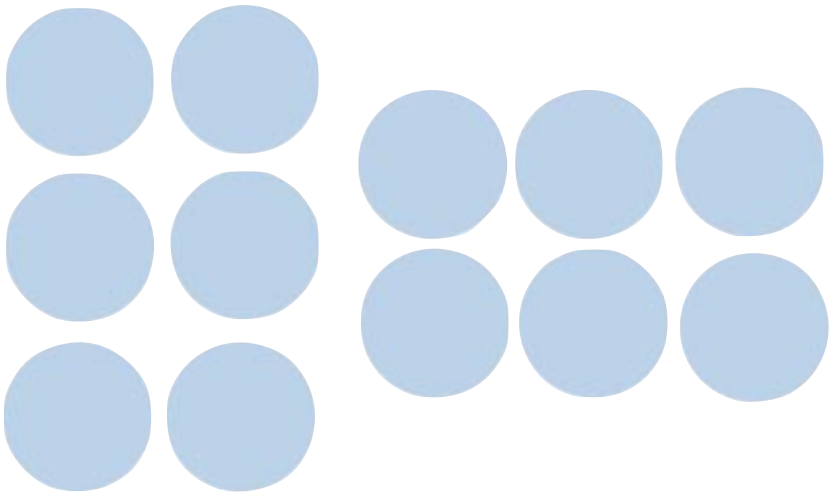
- Symmetrical images are perceived collectively, in spite of distance.



- We will most likely perceive the above image as two black symmetrical vessels against a light background (rather than three asymmetrical white objects against a black background)

Law of Proximity

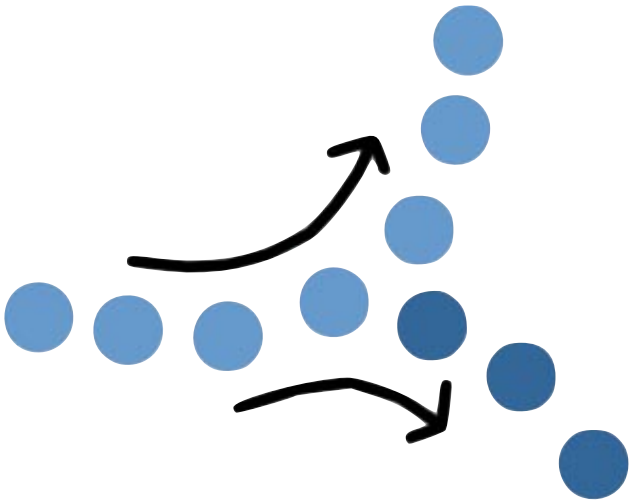
- Objects near each other tend to be grouped together.



- We will most likely see the circles on the left as grouped into two columns, and those on the right as grouped into two rows.

Law of Continuity

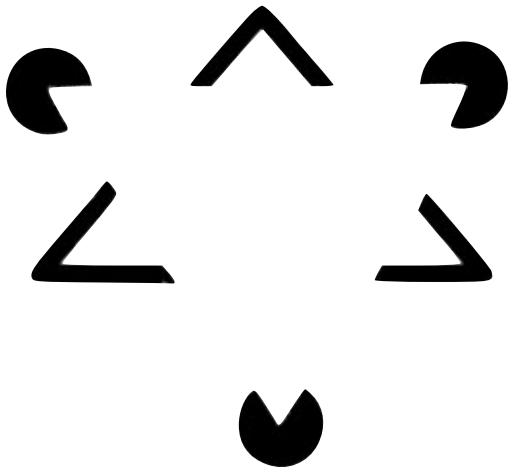
- Lines are seen as following the smoothest path.



- The top branch is seen as continuing the top segment of the line, rather than as an independent line.

Law of Closure

- Objects grouped together as seen as a whole. We tend to ignore gaps and complete lines.



- There are no circles or triangles in this picture. But the brain “fills in the gaps” with familiar figures.

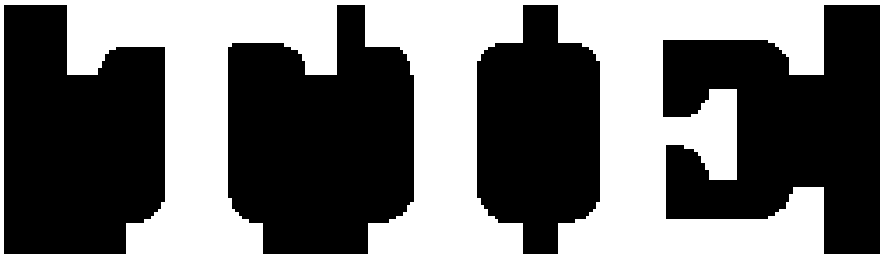
Law of Smallness

- Smaller areas tend to be seen as figures against a larger background.

- In the image above, “seeing” the vase is easier when it is smaller.

Law of Surroundedness

- Areas which are surrounded by others tend to be seen as figures.



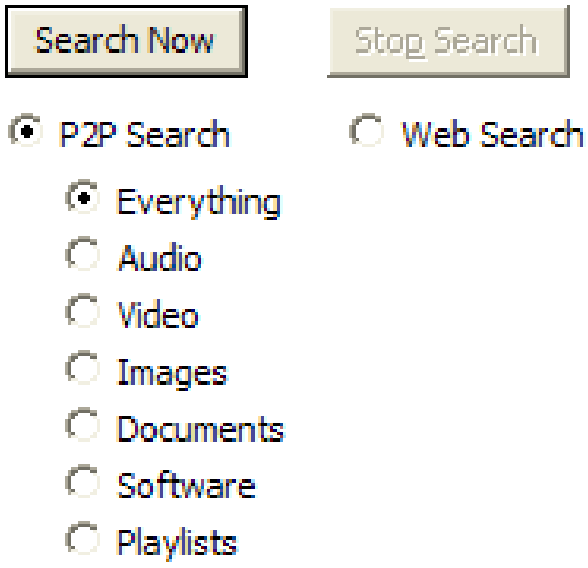
- Most people will see the word “TIE” in white letters against a black background.

Exploiting Gestalt Principles in Design

- Symmetry:
  - Try to lay out visual objects symmetrically
- Proximity and surroundedness
  - Place items that are related close to one another.
  - Use white space to distinguish between different groups.
- Similarity:
  - Use similar features on related items

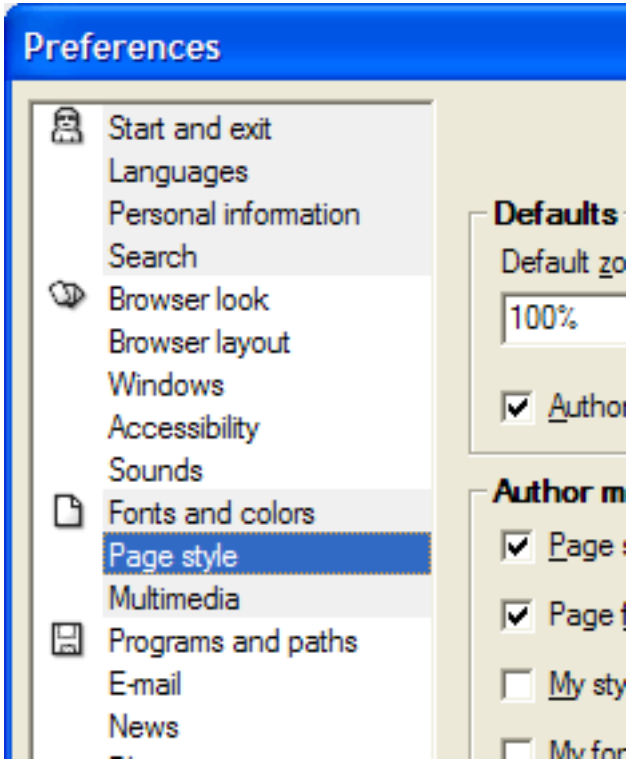
Gestalt Principles at Work: Proximity

Gestalt Principles at Work: Proximity



From Kazaa. Obviously those options are for “P2P Search”, and not for “Web Search”.

Gestalt Principles at Work: Similarity



From the Opera browser. Options with the same background color are "tied together" automatically by the user.

Gestalt Principles at Work: Closure

Iconic Input: Textual

- Three stages of reading:
  - Visual pattern perceived
  - Decoded using internal representation of language.
  - Interpreted using knowledge of syntax, semantics and pragmatics.

Textual Decoding and Interpretation

- Syntactic
  - Language Structure (i.e. “Grammar”)
  - E.g. Sentence = Noun_Phrase Verb_Phrase
- Semantic
  - The literal meanings of words and sentences, regardless of context
  - E.g. Have you been out in the rain? *No, I jumped into the lake.*
- Pragmatic
  - Meaning with regard to the individuals involved and the context of the situation

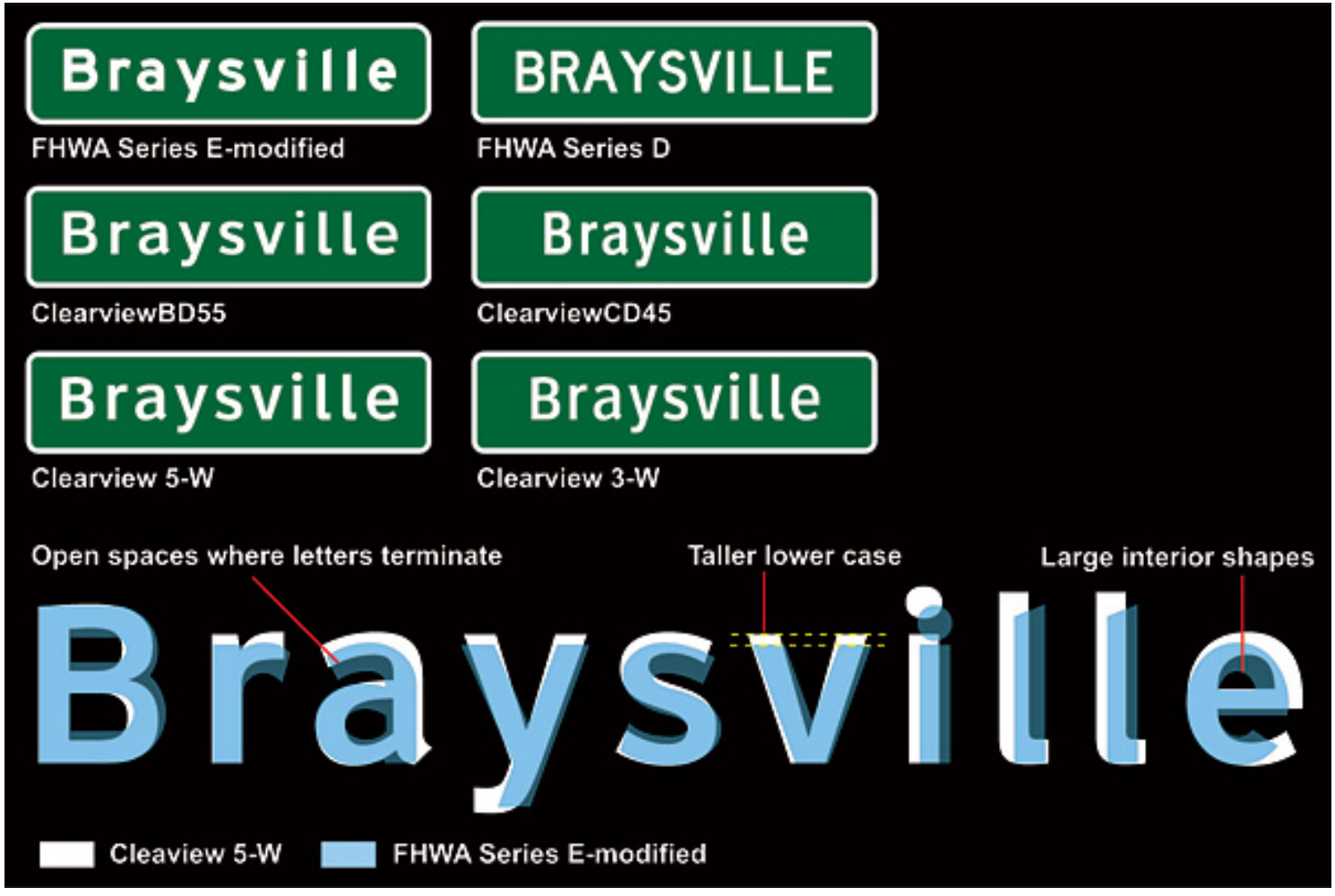
Textual Perception

- Reading involves saccades (scanning) and fixations.
  - Perception and comprehension takes place only during fixations.
  - Word shape (and subsequently, fonts) is important to recognition

Word Shape and Readability

Graphic from New York Times

Fonts and Readability



Graphic from New York Times

Fonts and readability

Graphic from New York Times

Echoic Input (Hearing)

- Provides information about environment: Distances, directions, objects, etc
- Physical apparatus:
  - Outer ear: Protects inner and amplifies sound
  - Middle ear: transmits sound waves as vibrations to inner ear
  - Inner ear: chemical transmitters are released and cause impulses in auditory nerve.

Echoic Input: Perception

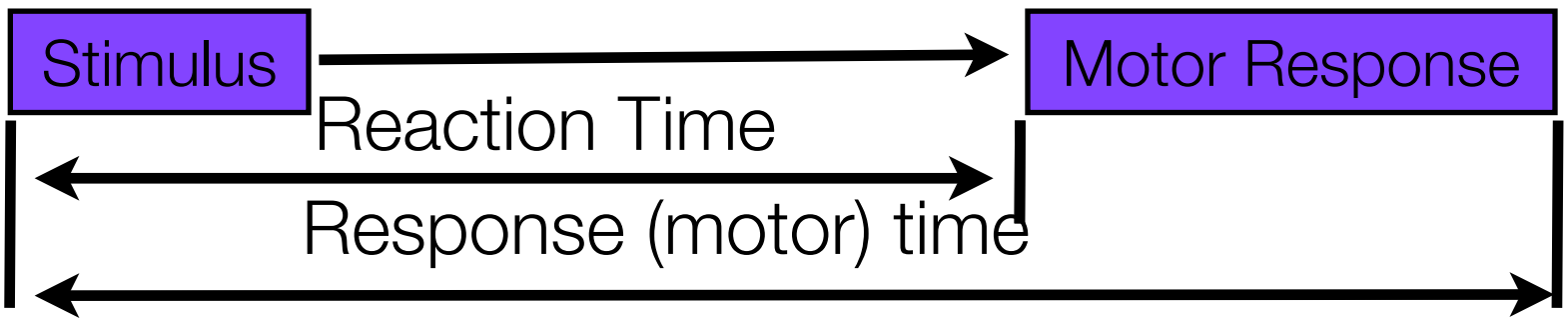
- Sound
  - Pitch: sound frequency
  - Loudness: amplitude
  - Timbre: type or quality
- Humans can hear frequencies from 20Hz to 15kHz, and are less accurate distinguishing high frequencies than low.
- We are selective in what we hear:
  - The auditory system filters sounds and we can attend to sounds over substantial background noise and concentrate on important information.
    - The cocktail party effect.

Haptic Input (touch)

- May be key sense for someone who is visually impaired.
- Stimulus received via receptors in the skin.
  - Thermoreceptors: heat and cold
  - Nociceptors: pain
  - Mechanoreceptors: pressure (instant, continuous)
- Some areas are more sensitive than others, like fingertips)

Motor Output

- Time taken to respond to stimulus (speed): reaction time + movement time
- Movement time is dependent on physical characteristics of subject: e.g. age, fitness level.
- Reaction time is dependent on stimulus type
  - Visual: ~200 ms
  - Auditory: ~150 ms
  - Pain: ~700 ms

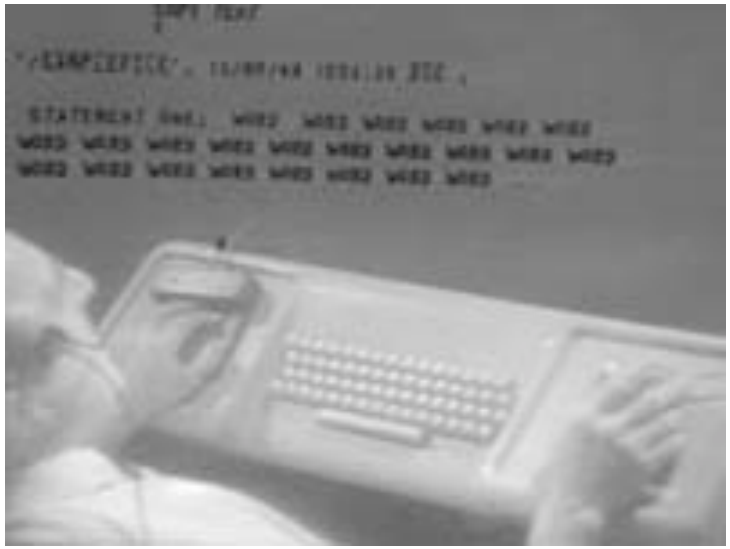


Motor Input: Accuracy

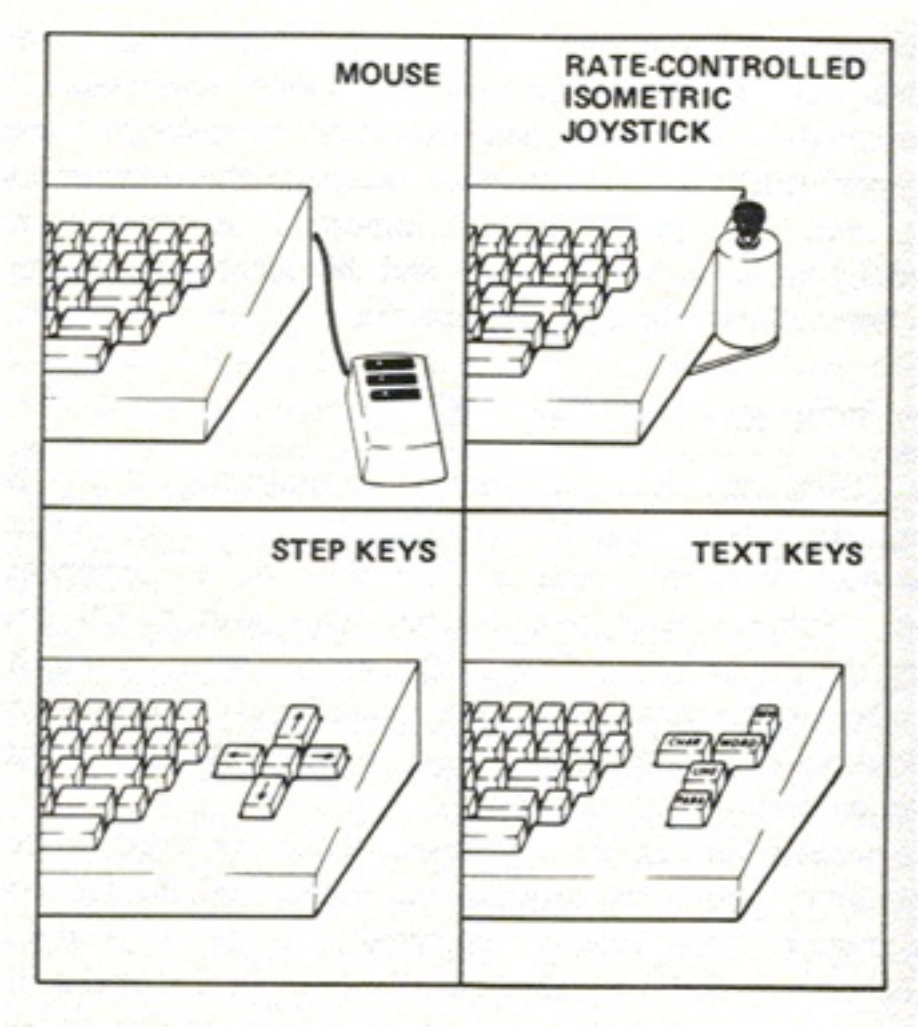
- Another measure of motor skill is accuracy
  - Whether speed of reaction results in reduce accuracy depends on the task and the user

Motor Input on Computers: The First Mouse (Engelbart & English, about 1964)

Which is fastest?

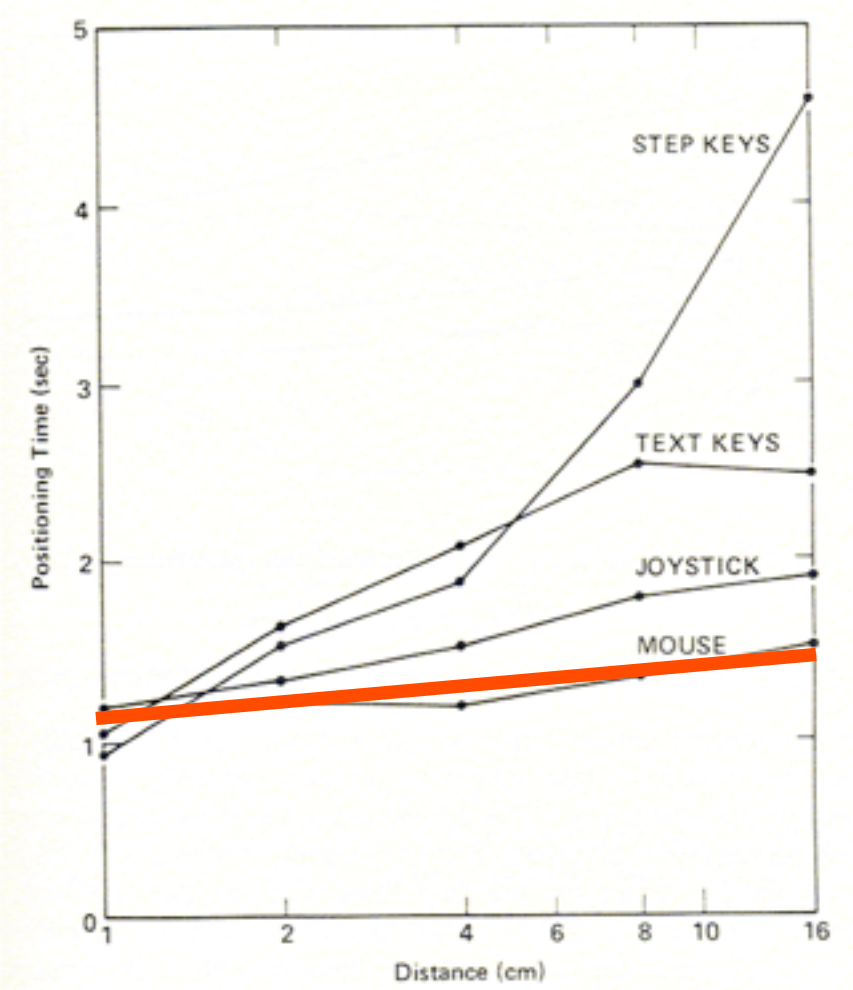


Engelbart



Source: Card, Stu. Lecture on Human Information Interaction. Stanford, 2007.

Experiment: Mice are fastest



Source: Card, Stu. Lecture on Human Information Interaction. Stanford, 2007.

Fitts' Law: Speed and Accuracy

- Established in 1954 by Prof. Paul Morris Fitts
  - Lieutenant-colonel of the US Air Force
  - Head of the US Air Force Human Engineering Division
  - Later professor at Ohio State U and U of Michigan
- Probably the *only* widely accepted law in HCI.

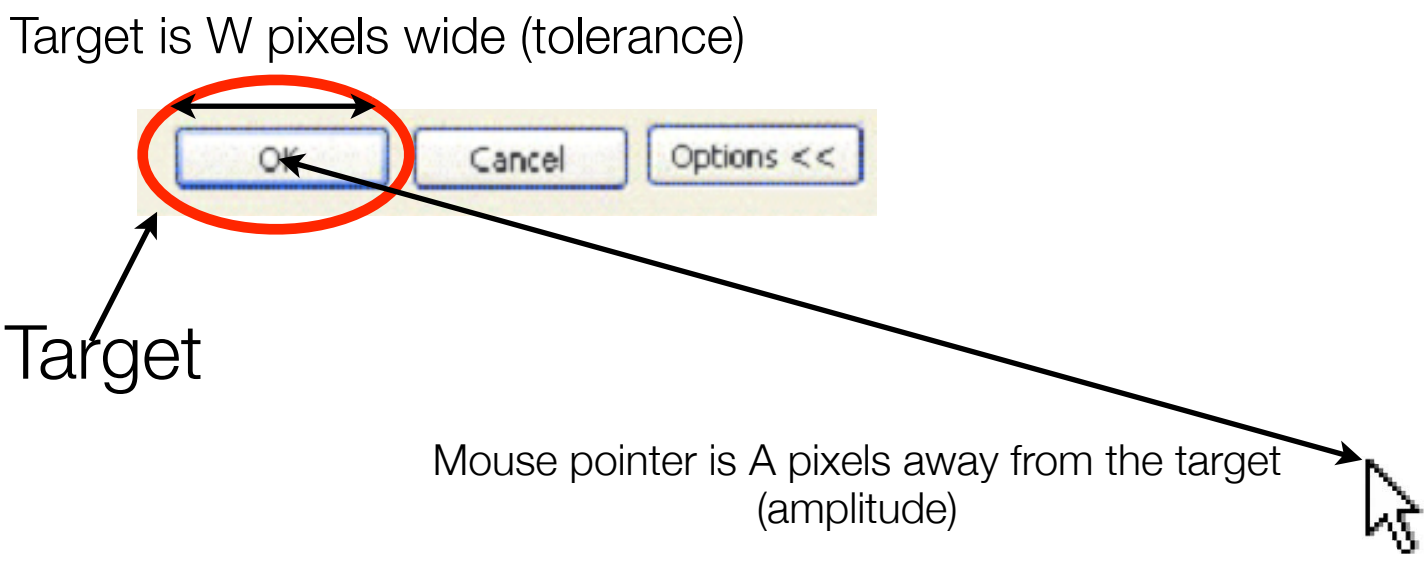
Fitts' Law

- A model of human movement
- Predicts the time required to rapidly move from a starting position to a final target area
- Models the act of pointing
  - In the real world (e.g., with a hand or finger, or even eye tracking)
  - On computers (e.g., with a mouse or a joystick)

Fitts' Law

- Model of human movement.
  - Based on Shannon's theory of information systems.
  - Human motor system modeled as channel used to transmit information
    - Task has a certain difficulty (measured by number of bits).
    - Channel has a certain bandwidth (measured by bits/second).

The Task



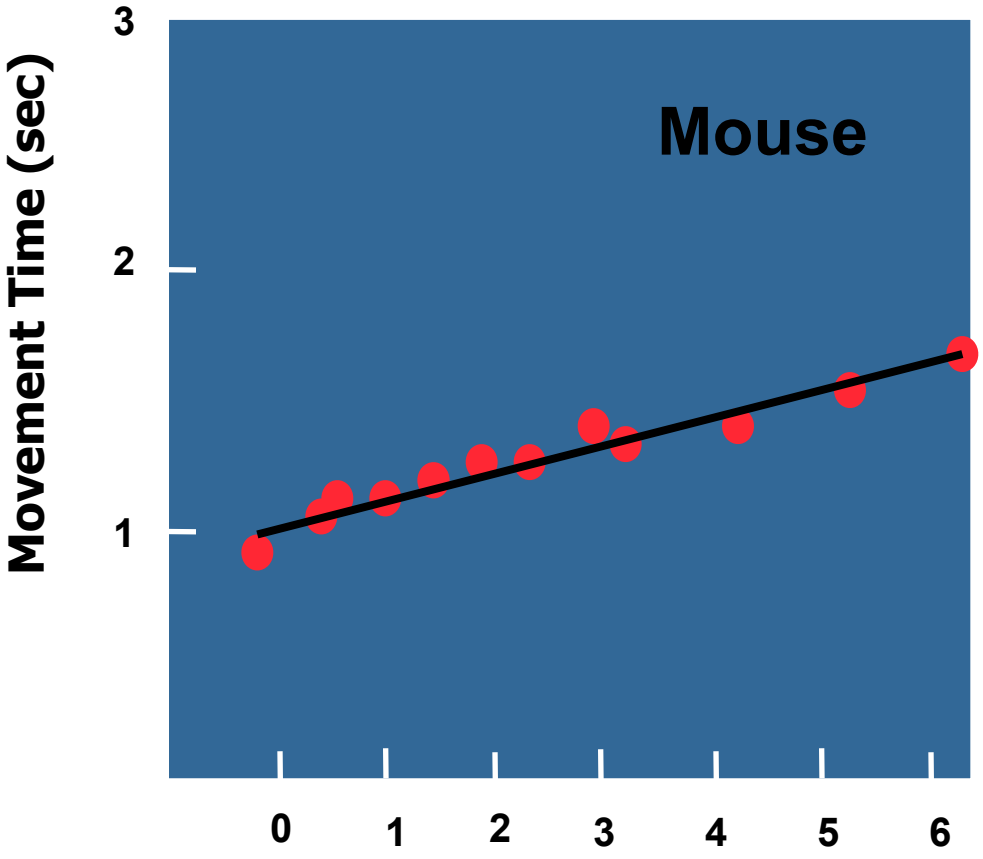
Index of Difficulty, or I_D, gives the *difficulty* of the task (in *bits*)
Depends both on A and W.

Index of Difficulty

- Movement to target requires more time when
  - Distance to target (A) increases
  - Error Tolerance of target (W) decreases
  - That is, when target is farther away or of smaller size
- Movement to target requires less time when
  - Distance to target (A) decreases
  - Error Tolerance of target (W) increases
  - That is, with a closer target or one of larger size.
- Originally: $I_D = \log_2(2A/W)$
- Shannon-style reformulation:
  - $I_D = \log_2(A/W + 1)$, or $I_D = \log_2(A/W + 0.5)$ (when $I_D < 3$ bits)

The Channel

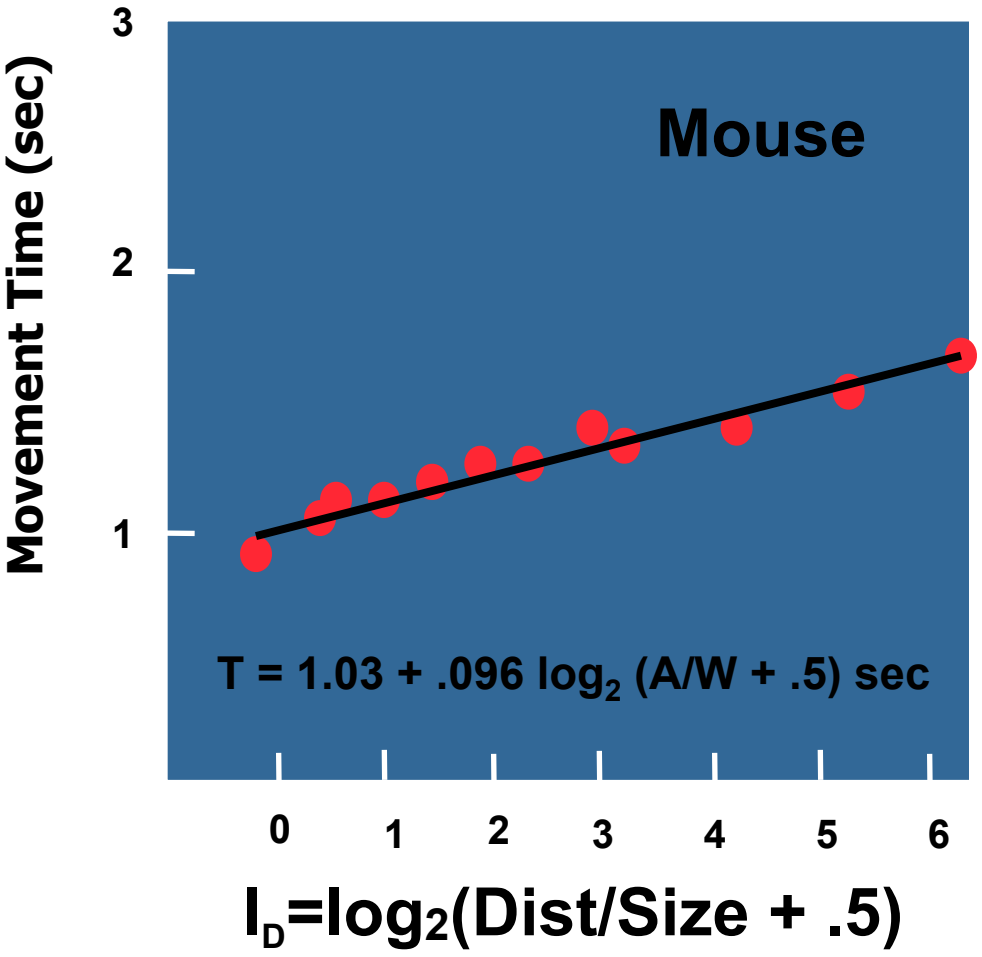
- The channel is characteristic of the *modality* of the interaction



Putting Everything Together

- Time T_{pos} needed to complete the task is linear with respect to I_D.
    - $T_{pos} = a + b*I_D$
- *a* and *b* are empirically determined through experimental data.
- *a* = y-intercept
- *b* = slope (bits/sec)
- T_{pos} = time taken (seconds)

Evaluating a device



Source: Card, Stu. Lecture on Human Information Interaction. Stanford, 2007.

The index of difficulty depends on the task, not the device.

The regression constants *a* and *b* determine how fast the task is done (or how fast the device processes the task).

The *index of performance* IP (or *throughput* TP) defines how quickly the task can be finished.

IP = 1/b

Aside: the mouse is very close to being an optimal device (harder to do better than the mouse).

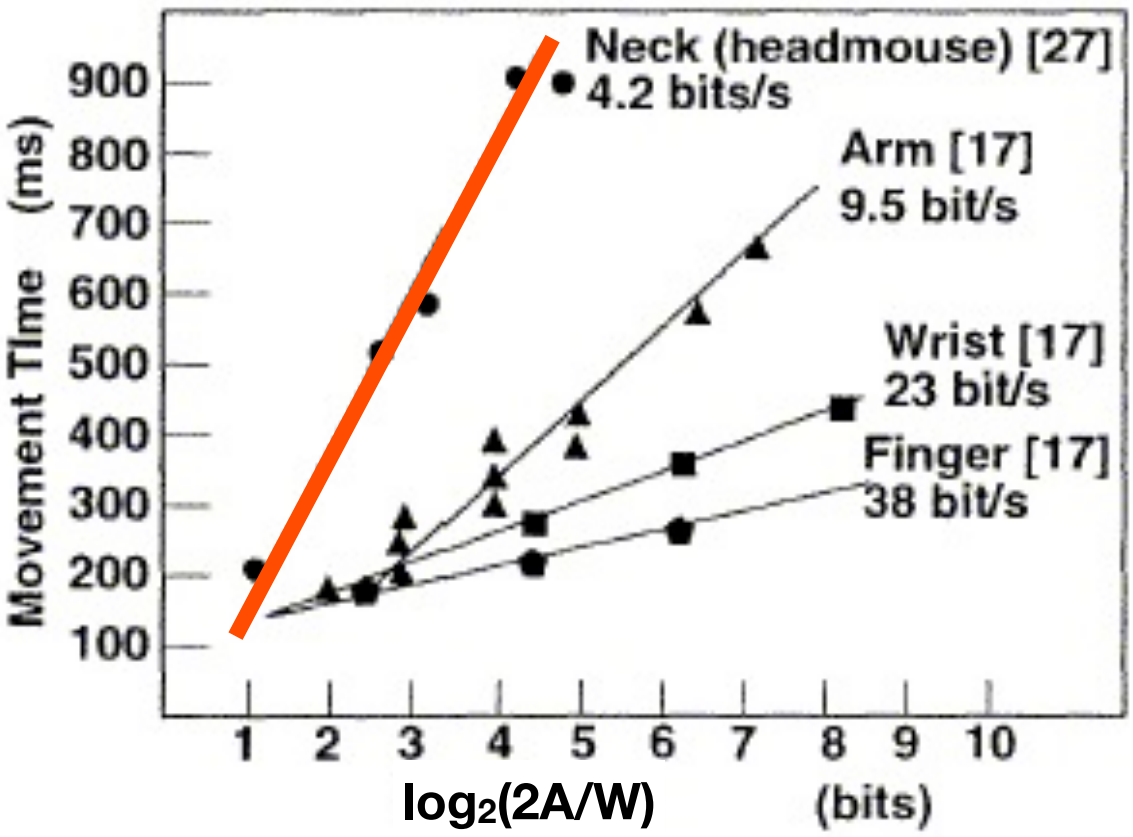
Adapting the Model

- Does this work for other devices?

Headmouse

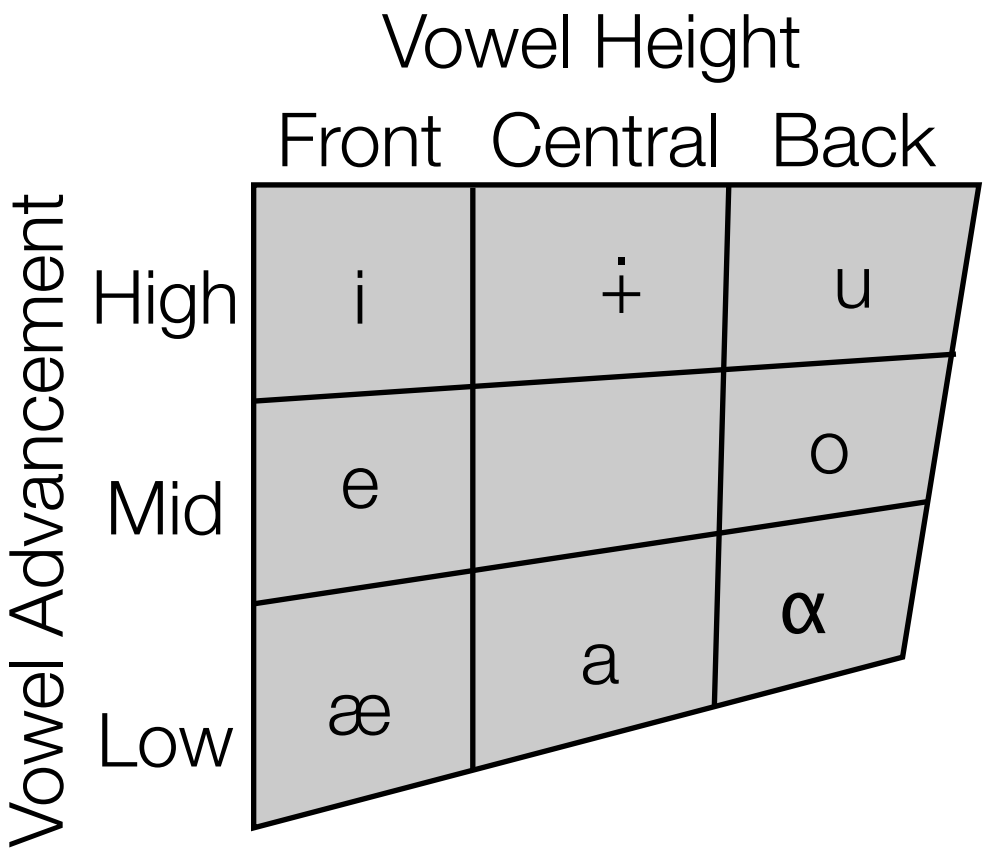
Source: Card, Stu. Lecture on Human Information Interaction. Stanford, 2007.

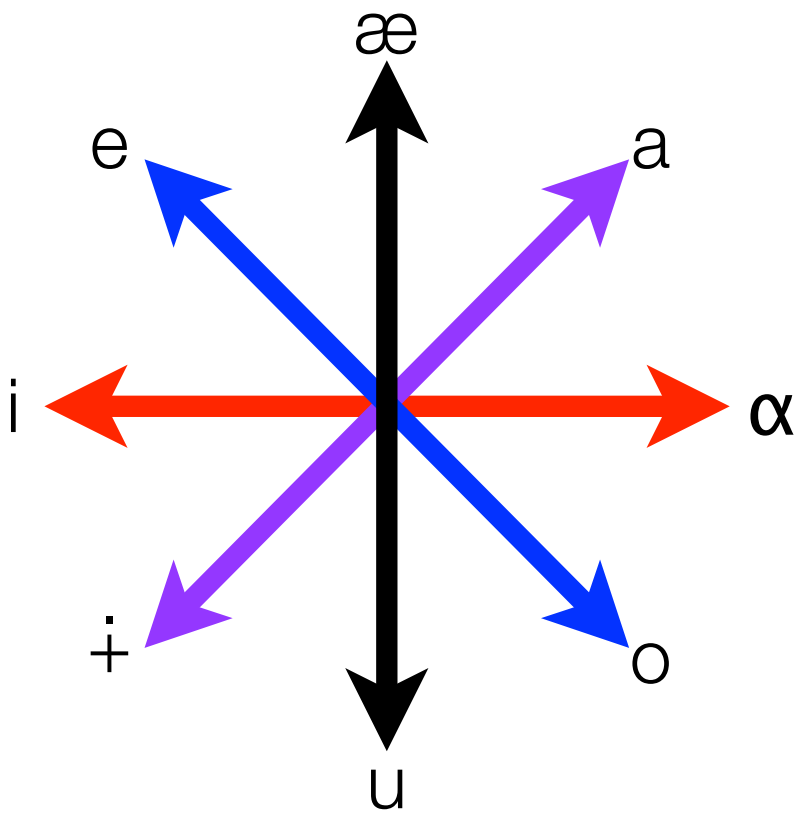
Also follows Fitts' Law



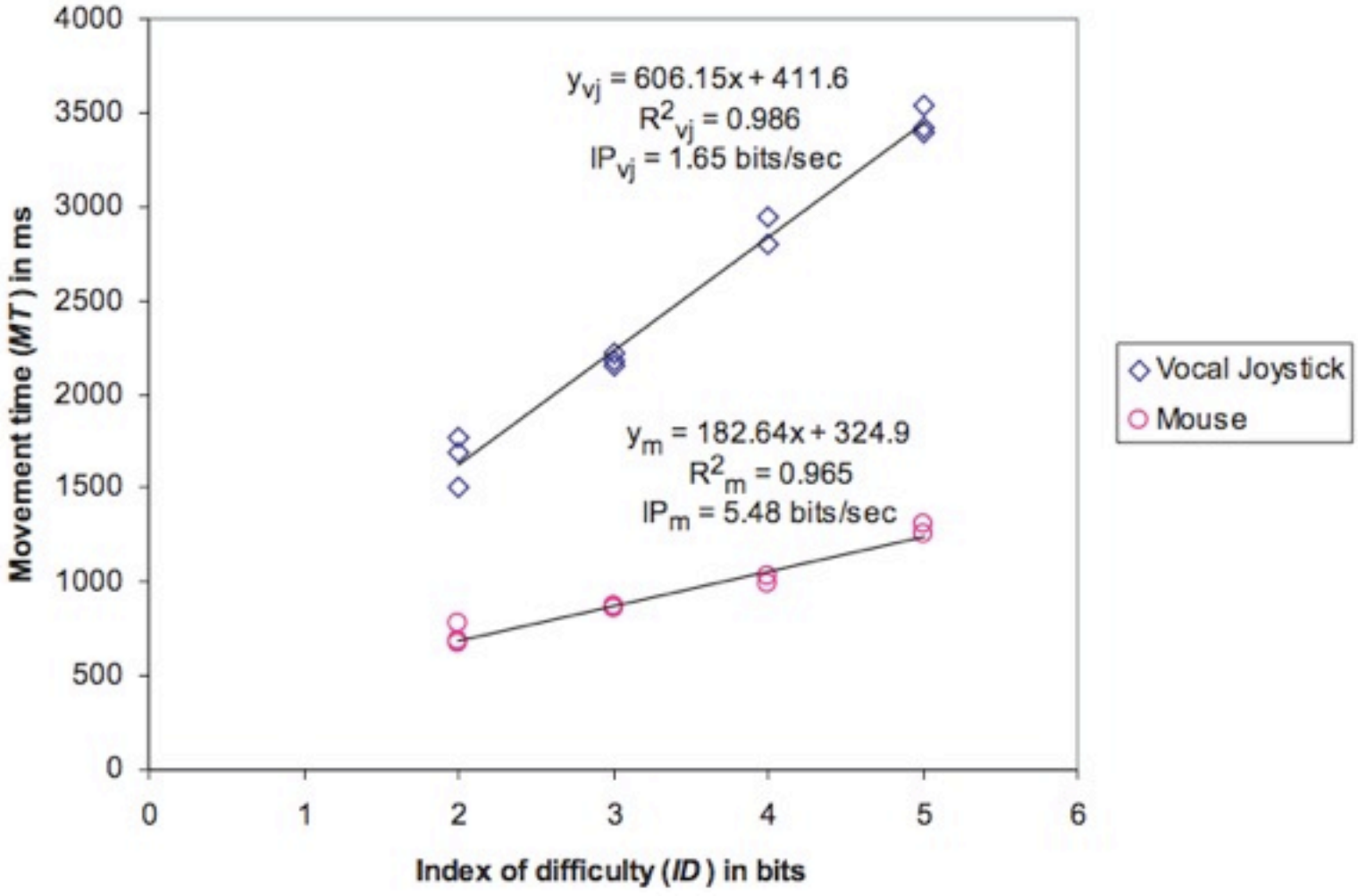
Source: Card, Stu. Lecture on Human Information Interaction. Stanford, 2007.

Vocal Joystick (Harada et al., 2006)



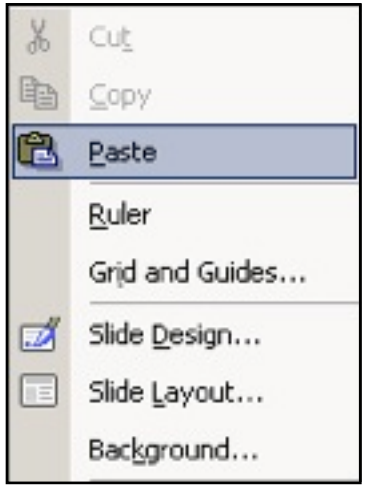


Also follows Fitts' Law



Exercise

- Calculate the time needed to reach a target that is:
  - 80 pixels wide
  - 200 pixels to the left of the cursor's current location.
  - $T_{pos} \sim= 590 + 230 \log_2(A/W+1)$



Adapting the Model

- How about vertical or diagonal movement (2 dimensions)?
- Or targets that are taller than they are wide?

More than 1 dimension

- Empirical studies show that tolerance, not direction, is the deciding factor.
- Measure tolerance by:
  - Area of target
  - Perimeter of target
  - Smaller of width and height
  - Length of line between center of target and target boundary (along line of approach)

Exercise

- Calculate the time needed to reach a target that is:
  - 400 pixels above and to the left of the cursor's current location.
  - 200 pixels long and 40 pixels wide.
  - $T_{pos} \sim= 590 + 230 \log_2(A/W+1)$

Implications of Fitts' Law

- Smaller targets that are farther away will take longer to point at.
- Applies to using mouse, finger, pen, stylus, even to focusing on it
- Group targets together if they are likely to be used in succession.
- Ensure targets are large enough for easy and quick pointing.

Memory

- We have three types of memory functions:

Sensory Memories

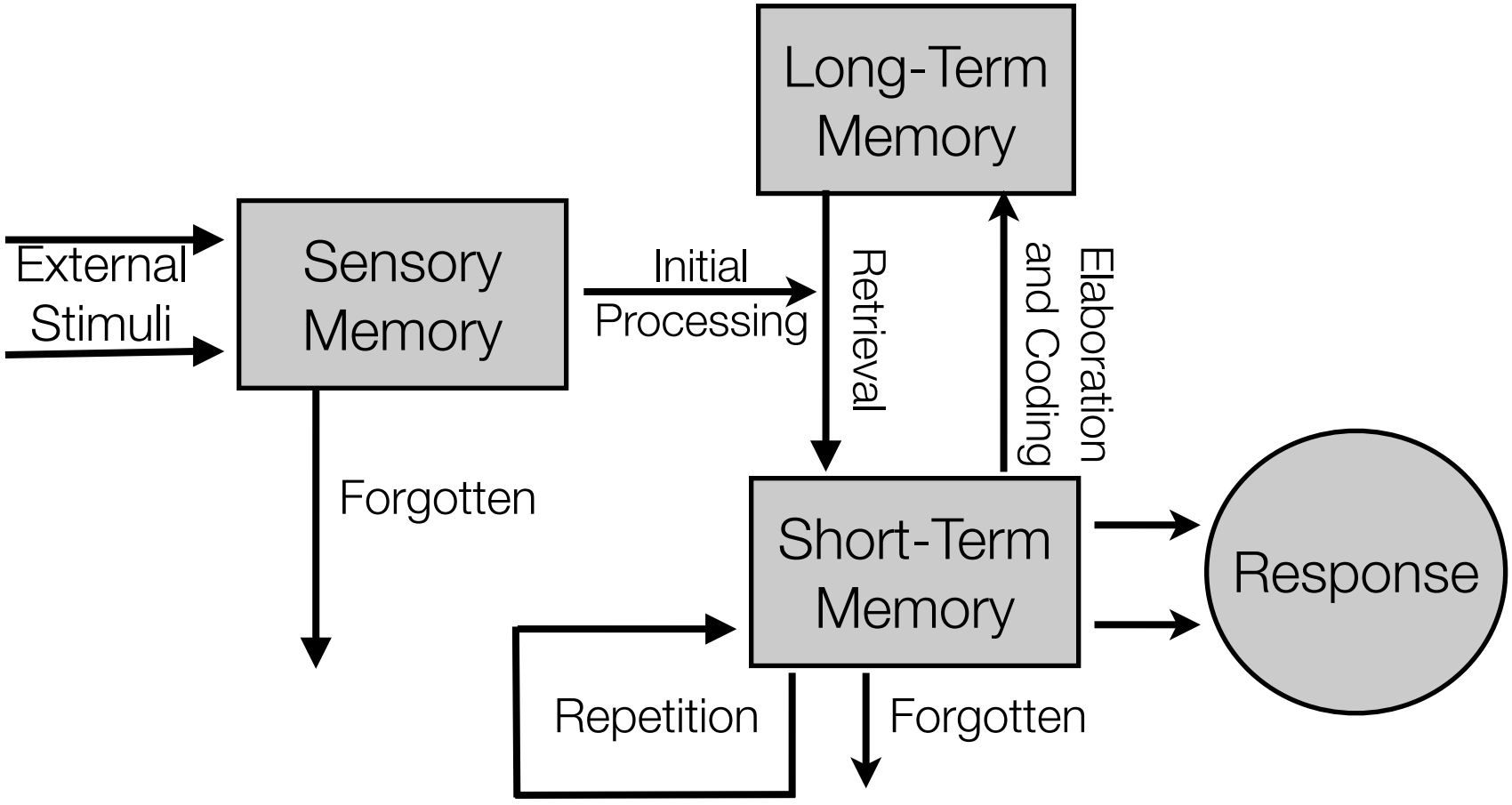
↓ **Attention**

Short-term Memory
(or working memory)

↓ **Rehearsal**

Long-term Memory

- Selection of stimuli governed by level of interest or need.


Long-Term Memory
External Stimuli
Sensory Memory
Initial Processing
Retrieval
Elaboration and Coding
Forgotten
Short-Term Memory
Response
Repetition
Forgotten

Sensory Memory

- Buffers for stimuli
  - Iconic: Visual
  - Echoic: aural
  - Haptic: touch
- Constantly overwritten
- Information passes from sensory to STM by attention
- Selection of stimuli governed by level of arousal

Short-Term Memory (STM)

- Also known as Working Memory
- Scratch pad for temporary recall
- Transient
  - Rapid Decay: 200ms
  - Rapid Access: 70ms
- Limited capacity: 7 ± 2
- Subject to interference
- Forgets if not rehearsed

Short Term Memory

- A typical experiment
  - Subjects study list of nonsense items
  - They then try to recall the list over varying periods of time
  - No rehearsal
  - Recall deteriorated as a function of the duration of the retention interval

- (STM Experiment)

Short Term Memory

- Primary Effect: Recall of the first few items is better than recall of later items
- Recency Effect: Recall of most recently seen items is better than recall of earlier items
- Some evidence for several elements of STM -- articulatory channel, visual channel etc -- interference on different channels does not impair recall.

Long-Term Memory

- Depository for all our knowledge
- Slow access: 0.1 second
- Slow decay, if any
- Huge or unlimited capacity
- Two types:
  - Episodic: serial memory of events
  - Semantic: structured memory of facts, concepts, skills

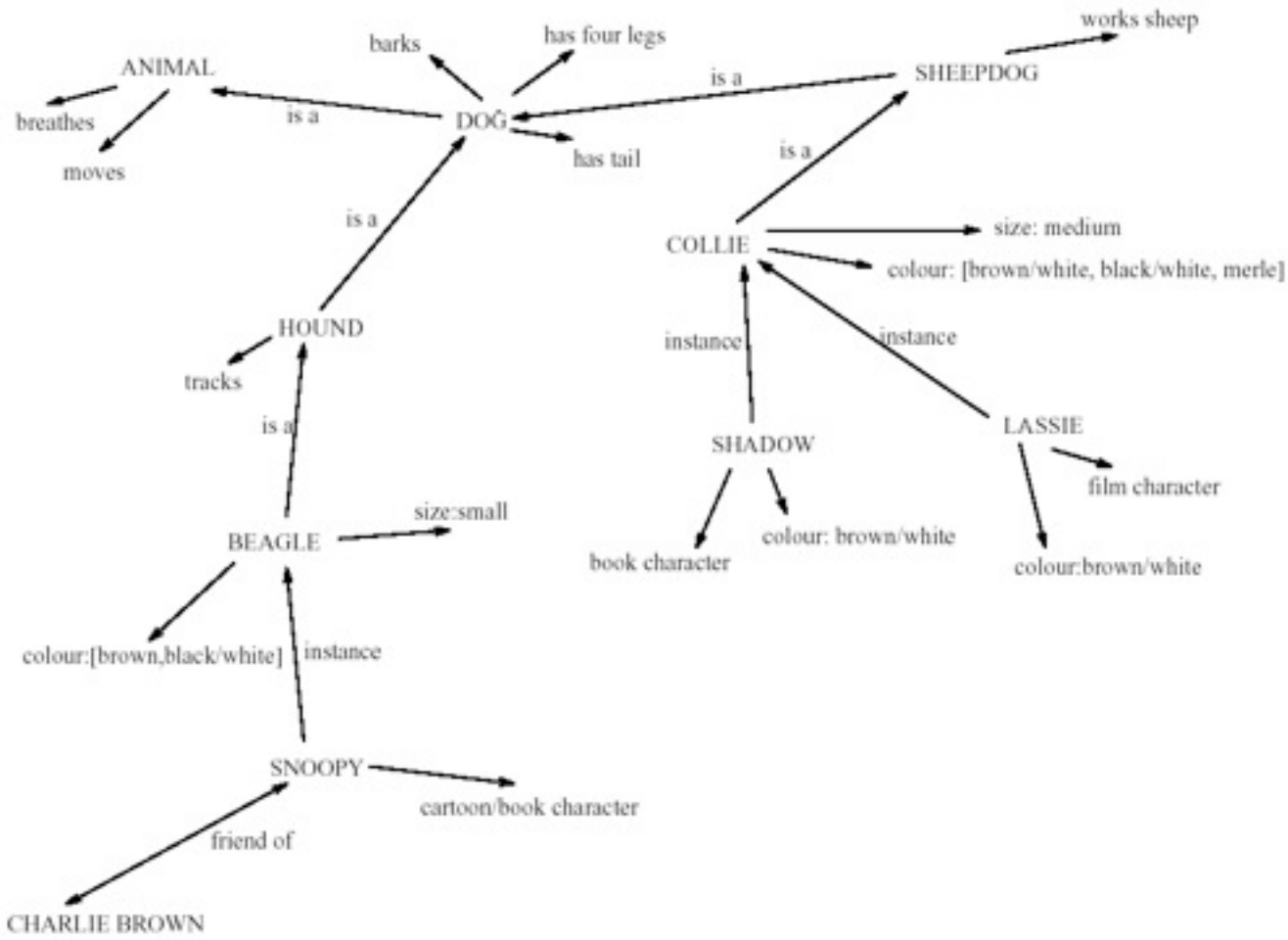
Long-Term Memory

- Information in semantic LTM derived from episodic LTM
- Semantic Memory Structure
  - Provides access to information
  - Represents relationships between bits of information
  - Supports inference

LTM Models

- Semantic memory structure
  - Provides access to information
  - Represents relationships between bits of information
  - Supports inference
- Model: Semantic Network
  - Inheritance: Child nodes inherit properties of parent nodes
  - Relationships between bits of information explicit
  - Supports inference through inheritance

Semantic Network



Models of LTM — Frames

- Information organized in data structures
- Slots in structure instantiated with values for instance of data
- Type–subtype relationships

DOG

Fixed
legs: 4

Default
diet: carnivorous
sound: bark

Variable
size:
colour

COLLIE

Fixed
breed of: DOG
type: sheepdog

Default
size: 65 cm

Variable
colour

Models of LTM — Scripts

- Model of stereotypical information required to interpret situation
- Script has elements that can be instantiated with values for context

Script for a visit to the vet

Entry conditions:	*dog ill* *vet open* *owner has money*
Result:	*dog better* *owner poorer* *vet richer*
Props:	*examination table* *medicine* *instruments*
Roles:	*vet examines* *diagnoses* *treats* *owner brings dog in* *pays* *takes dog out*
Scenes:	*arriving at reception* *waiting in room* *examination* *paying*
Tracks:	*dog needs medicine* *dog needs operation*

Models of LTM — Production Rules

- Representation of procedural knowledge.
- Condition/action rules
  - if condition is matched
    then use rule to determine action.

IF dog is wagging tail
THEN pat dog

IF dog is growling
THEN run away

LTM Processes: Storage

- Information moves from STM to LTM by *rehearsal*
- Amount retained proportional to rehearsal time: Total Time Hypothesis
- Optimized by spreading learning over time distribution of practice effect
- Structure, meaning and familiarity make information easier to remember.

LTM Processes: Forgetting

- Decay:
  - Information is lost gradually but very slowly
- Interference
  - New information replaces old: retroactive interference
  - Old may interfere with new: proactive inhibition

So may not forget at all. Memory is selective...
... Affected by emotion, so can “choose to forget”

LTM Processes: Retrieval

- Recall
  - Information reproduced from memory
  - Can be assisted by cues: e.g. Categories, imagery
- Recognition
  - Information gives knowledge that it has been seen before
  - Less complex than recall -- information is cue.

Thinking

- Reasoning
- Deduction, induction, abduction
- Problem solving

Reasoning: Deduction

- Derive logically necessary conclusion from given premises.

  e.g. If it is Friday then she will go to work

  It is Friday

  Therefore she will go to work.

- Note: Just because a conclusion is logical does not mean it is true:

  e.g. If it is raining then the ground is dry

  It is raining

  Therefore the ground is dry

Reasoning: Deduction (cont.)

- When truth and logical validity clash …

e.g. Some people are babies

Some babies cry

Inference - Some people cry

Correct?

- People use world knowledge to resolve.

Reasoning: Induction

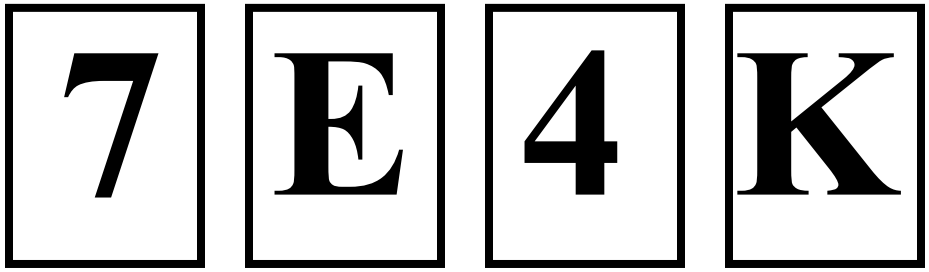
- Generalize from cases seen to cases unseen

  e.g. all elephants we have seen have trunks
  therefore all elephants have trunks.

- Induction is unreliable:
  - It can only prove that something is false
  - But in general, it does prove to be useful.
- Humans are not good at using negative evidence

  e.g. Wason's cards.

Wason's Cards



Given: 4 cards, a letter on one side, a number on the other.

Hypothesis: If a card has a vowel on one side it has an even number on the other

- Is this true?
- How many cards do you need to turn over to find out?
- Which cards?

Reasoning: Abduction

- Reasoning from event to cause

  e.g. Sam drives fast when drunk.

  If I see Sam driving fast, assume he is drunk.

- Abduction is often unreliable:
  - Can lead to false explanations

Emotion

- Various theories of how emotion works
  - James-Lange: emotion is our interpretation of a physiological response to a stimuli
  - Cannon: emotion is a psychological response to a stimuli
  - Schacter-Singer: emotion is the result of our evaluation of our physiological responses, in the light of the whole situation we are in
- Emotion clearly involves both cognitive and physical responses to stimuli

Emotion

- The biological response to physical stimuli is called *affect*
- Affect influences how we respond to situations
  - positive → creative problem solving
  - negative → narrow thinking
- Negative affect can make it harder to do even easy tasks; positive affect can make it easier to do difficult tasks. (Don Norman)

Emotion (cont.)

- Implications for interface design
  - Stress will increase the difficulty of problem solving
  - Relaxed users will be more forgiving of shortcomings in design
  - Aesthetically pleasing and rewarding interfaces will increase positive affect

Individual differences

- Long term
  - Gender, physical and intellectual abilities
- Short term
  - Effect of stress or fatigue
- Changing
  - Age

Ask yourself:
will design decision exclude section of user population?

Psychology and the Design of Interactive System

- What does this have to do with HCI?
- There are some direct applications
  - e.g. blue acuity is poor
    $\Rightarrow$ blue should not be used for important detail
- However, correct application generally requires understanding of context in psychology, and an understanding of particular experimental conditions
- A lot of knowledge has been captured in design guidelines and cognitive models