Lecture 4

The Human

UNIVERSITY OF AUCKLAND
SOFTENG 350
Prof. Robert Amor

Notes from

The UX Book Chapter 32.3

Dix et al Chapter 1

Preece Chapter 3

www.id-book.com

Robot irony

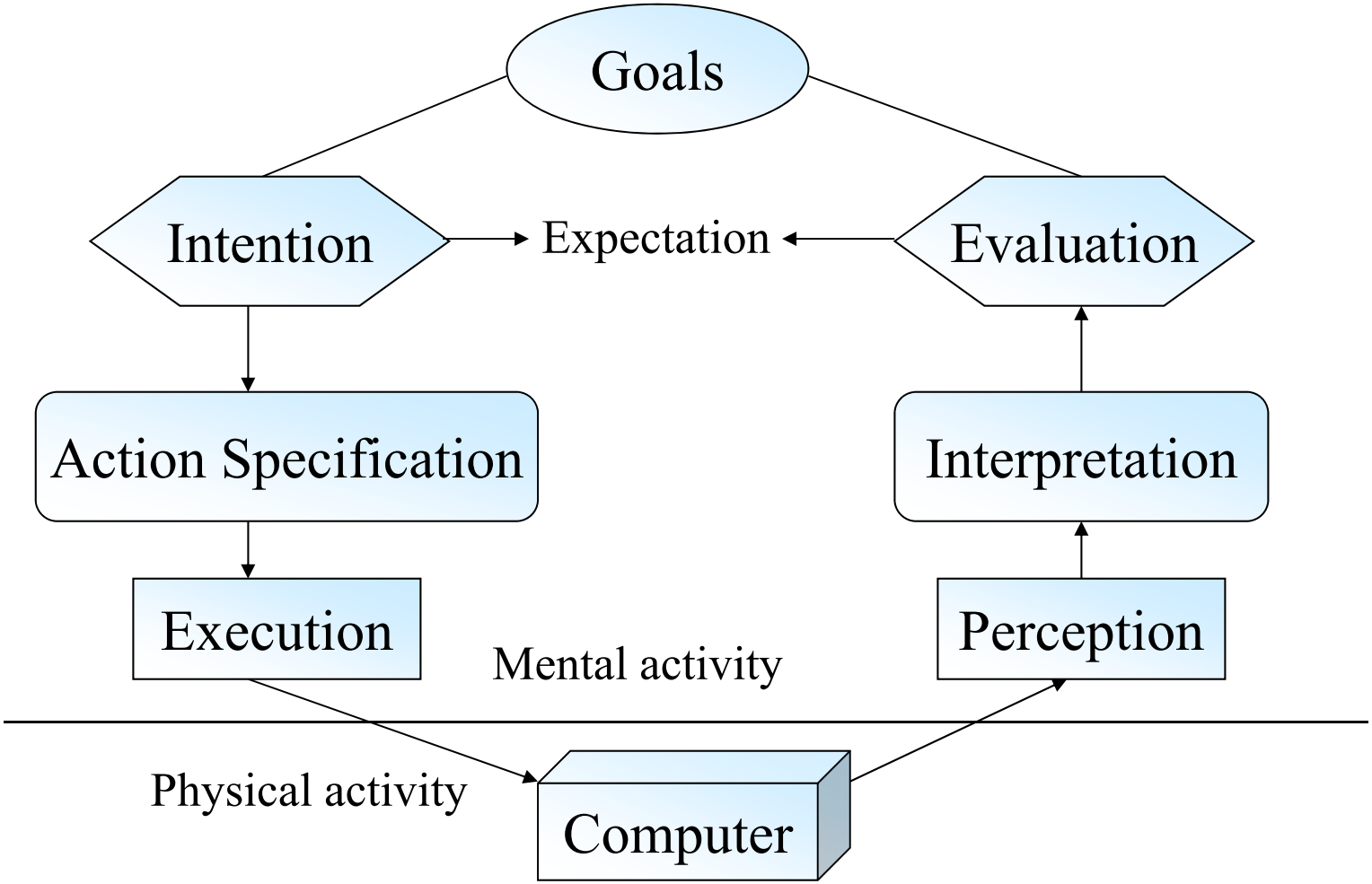
Learning objectives

- To be aware of the characterisation of a 'model human processor'
- To understand the impact of human capabilities on computer interaction
- To understand how our memory, cognition and senses operate

Design for human capabilities

- Human processor
  - Input
    - Senses
  - Storage
    - memory
  - Process
    - Cognition
      - Knowledge
      - Skills
      - Reasoning
  - Output
    - Actions
  - Why do we need to know this for HCI?

The Human Information Processor



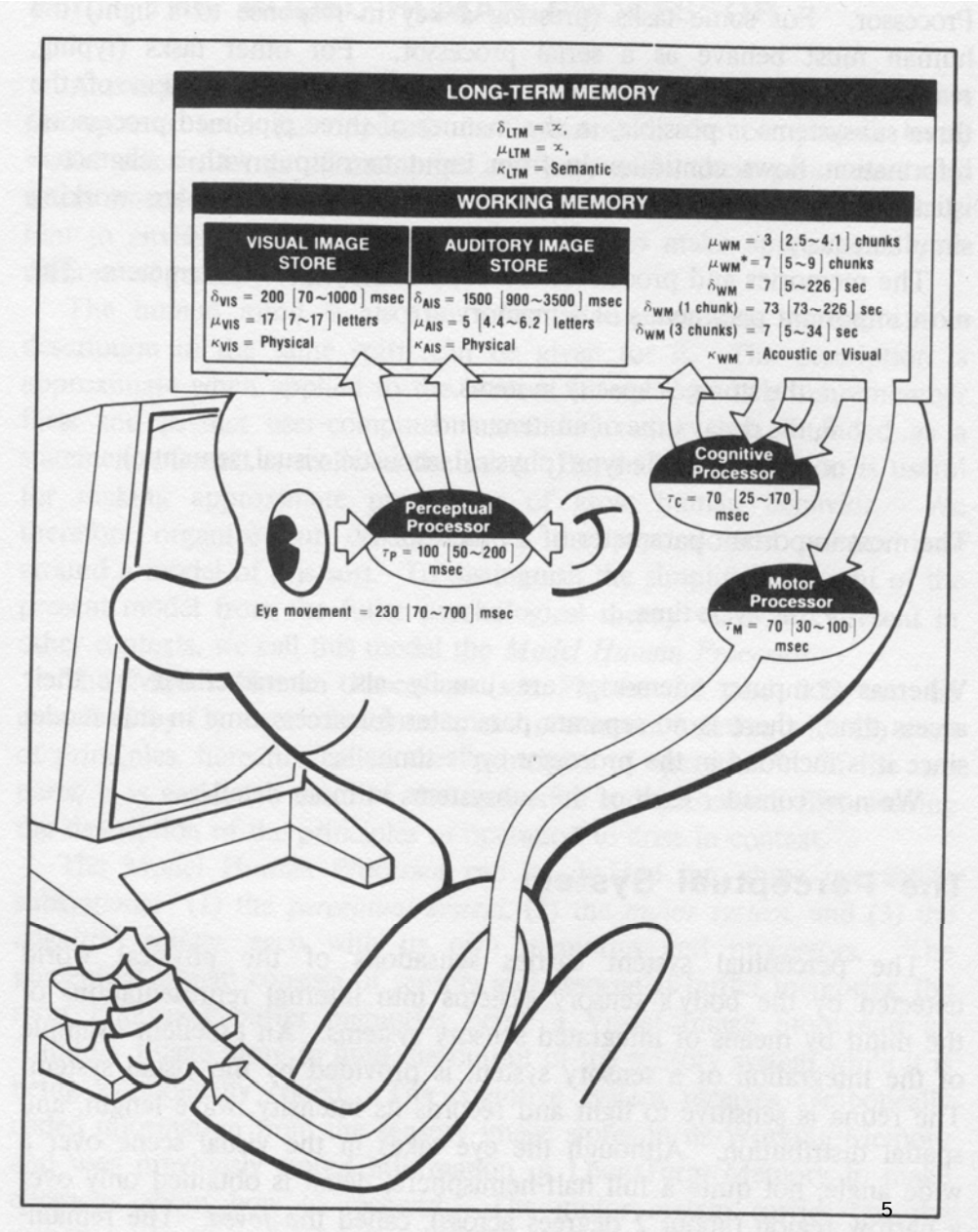
The Model Human Processor

μ – Storage capacity
δ – Decay time (half life)
κ – Code type (visual/auditory)

δ_{VIS} = 200[90-1000] msec
δ_{AIS} = 1500[900-3500] msec

μ_{VIS} = 17[7-17] letters
μ_{AIS} = 5[4.4-6.2] letters

Eye movement = 230[70-700] msec



Memory

There are three types of memory function:

Sensory memories

Attention

Short-term memory or working memory

Rehearsal

Long-term memory

Selection of stimuli governed by level of arousal.

Sensory memory

- Buffers for stimuli received through senses
  - iconic memory: visual stimuli
  - echoic memory: aural stimuli
  - haptic memory: tactile stimuli
- Examples
  - “sparkler” trail
  - stereo sound
- Continuously overwritten

Short-term memory (STM)

- Scratch-pad for temporary recall
  - Rapid access ~ 70ms
  - Duration ~ 30s
  - Limited capacity - 7± 2 chunks

- Some research suggests that programmers have better short-term memory than 'average' people
  - This means you will have better short-term memory than your users!

- Experience moves you from recognition to recall
  - Recognition can become a barrier (e.g., move to hot keys)

Memory Experiments

Memory Experiments

- Digit span – measure of a good memory
- Chunking (7 +/- 2 items in short term mem)
  - If you can chunk information then can remember more
- Primacy and recency effects
  - First items are rehearsed several times and retained
  - Final items are remembered as no decay yet

- Closure, users are relieved when they have completed a task and don't have to remember items any more (eg dialing number from phonebook)
  - Task context is 'popped' off the memory stack, reduces cognitive load

- Recall works best in a similar environment

- Extent that new material can be remembered depends upon its meaningfulness

Memory Experiments

Memory Experiments

Stacking

- Task context stacking
  - When a task is interrupted
  - Original context goes on a ‘stack’ in user working memory
  - User task stacks are small and don’t last long

Interesting trivia

- Experienced programmers
  - Use chunked techniques that they have used before to solve problems
  - They decompose the problem into bigger chunks than a novice programmer
  - Then apply known solution to each chunk
  - Takes about 10 years to build up repertoire

Long-term memory (LTM)

- Repository for all our knowledge
  - slow access ~ 1/10 second
  - slow decay, if any
  - huge or unlimited capacity

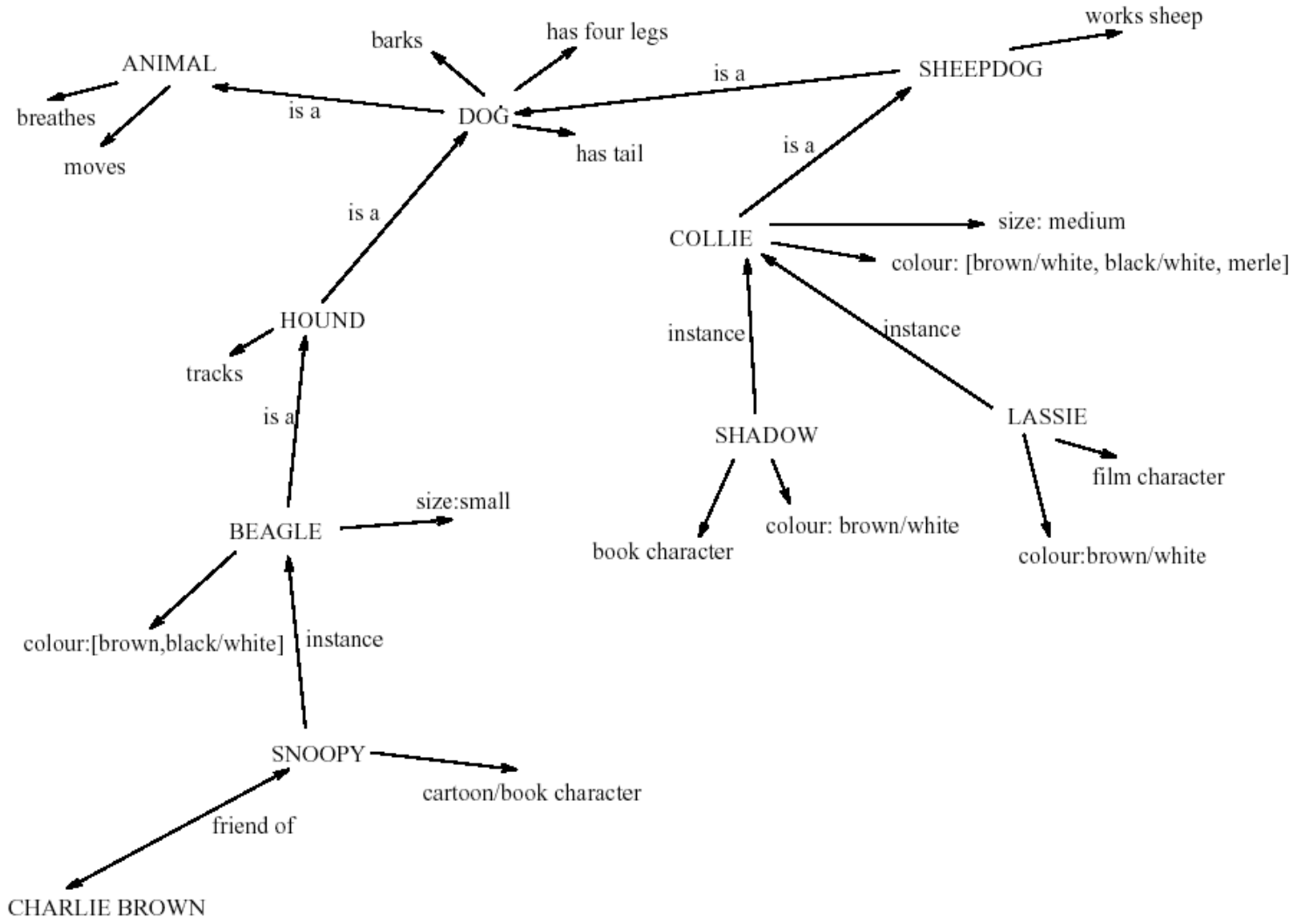
- Two main types
  - episodic – serial memory of events
  - semantic – structured memory of facts, concepts, skills

  semantic LTM derived from episodic LTM

Long-term memory (cont.)

- 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

LTM - semantic network



LTM - Forgetting

Decay

- information is lost gradually but very slowly

Interference

- new information replaces old: retroactive interference
- old may interfere with new: proactive interference

So may not forget at all, memory is selective …

… affected by emotion – can subconsciously `choose' to forget

LTM - 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

Processes

- Perceptual process
  - τ_P = 100[50-200] msec
  - Found with light pulses, sound clicks
    - E.g., for 0.1sec instant response, 20 frame/sec for film/TV
- Motor system
  - τ_M = 70[30-100] msec
- Working and Long-Term Memory
  - δ_{WM} = 7[5-226] sec
    - 1 chunk = 73 sec, 3 chunks = 7 sec
  - μ_{WM} = 7[5-9] chunks
  - $\delta_{LTM} = \infty$

Cognitive Processes

- Cognitive processor
  - τ_C = 70[25-170] msec

- Matching items to working memory (Cavanaugh 1972)

Digits	33[27-39] msec/item
Colors	38 msec/item
Letters	40[24-64] msec/item
Words	47[36-52] msec/item
Geometrical shapes	50 msec/item
Random forms	68[42-93] msec/item
Nonsense syllables	73 msec/item

Cognitive Processes

- Controlled vs. automatic cognitive processes
  - Automatic processes are not affected by the limited capacity of the brain and do not require attention
  - Controlled processes have limited capacity and require attention and conscious control
- Automatic processes are difficult to change once learnt, but are:
  - Fast
  - Demanding minimal attention
  - Unavailable to consciousness
  - Cognitive interference (e.g., Stroop effect – mismatched stimuli)

Stroop effect

Cognitive models are useful in:

- Initially constraining the design space.
- Answering specific design decisions.
- Estimating the total time for task performance with sufficient accuracy.
- To calculate the training time and guide the training documentation.
- Knowing which stages of activity take the longest time or produce most errors.

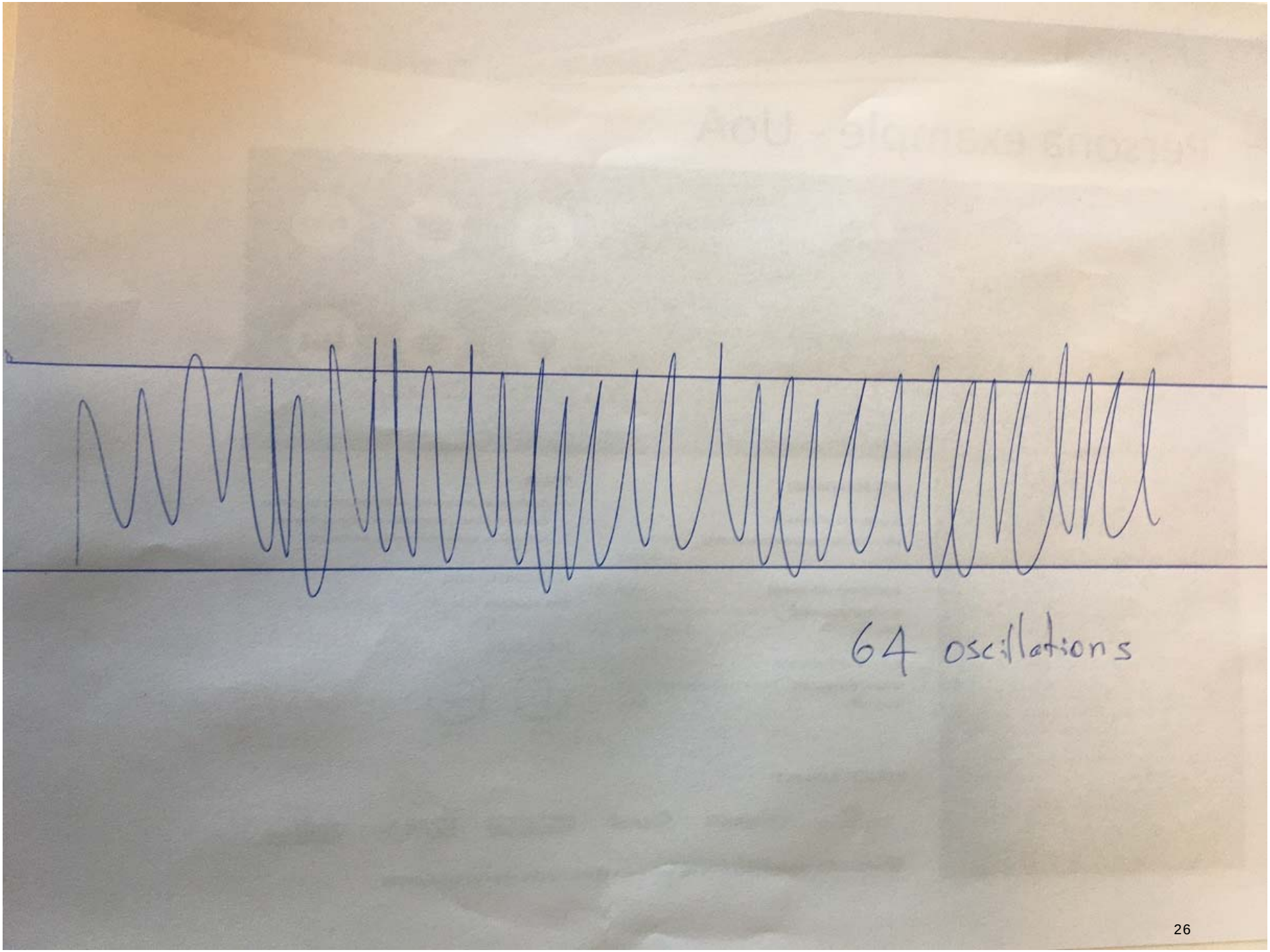
Example Calculations

- Reading text
  - = (60 sec/min)(0.23 sec/saccade)(5 saccade/word)
  - = 52 words/min

- Matching a symbol
  - Presented with 2 symbols one at a time. User presses a YES key if they match
  - = $\tau_P + 2\tau_C + \tau_M$
  - = 100 + 2*70 + 70
  - = 310[130-640] msec

Cycle times and task performance

- Draw an oscillating trace as fast as possible between the parallel lines for 5 seconds
- Drawing is limited by motor subsystem
  - Signal a reversal in direction every 70msec
    - Expect about 71 oscillations (5000/70)
- Model predicts how rapidly ends of travel are corrected
  - Perceptual system follows the trace, cognitive system decides a correction is needed, motor subsystem makes a correction
  - Sum of cycle time 100 + 70 + 70 = 240msec
    - Maximum number of corrections will be about 20 (5000/240)


64 oscillations

Multi-choice

Cognitive affordances which improve usability for novice and intermediate users can be an impediment for expert users. For example, learning a hot key rather than mouse clicks on menus. Which of Nielsen's heuristics needs to be violated to support this learning:

a) Flexibility and efficiency of use
b) Visibility of system status
c) Recognition rather than recall
d) User control and freedom

Senses

- Vision
- Hearing
- Touch (haptics)
- Smell
- Taste

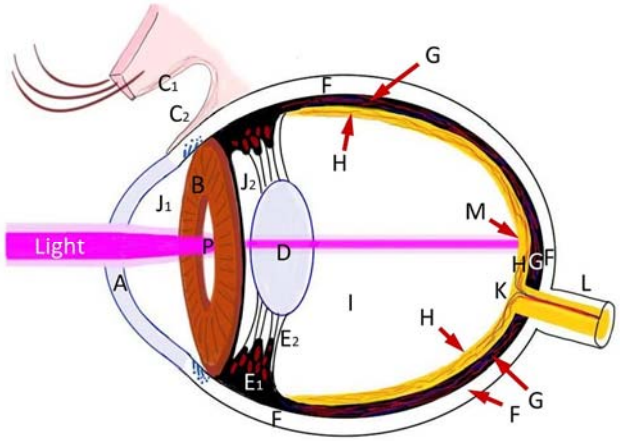
Vision

Two stages in vision

- Physical reception of stimulus
- Processing and interpretation of stimulus

The Eye - physical reception

- Mechanism for receiving light and transforming it into electrical energy
- Light reflects from objects
- Images are focused upside-down on retina
- Retina contains rods for low light vision and cones for colour vision
- Ganglion cells (brain!) detect pattern and movement



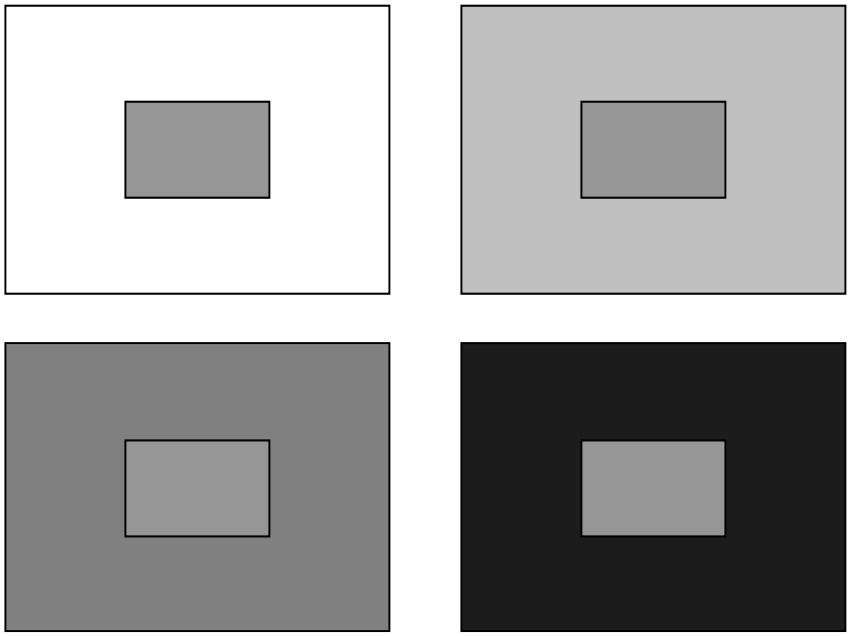
Interpreting the signal

- Size and depth
  - visual angle indicates how much of view object occupies (relates to size and distance from eye)
  - visual acuity is ability to perceive detail (limited)
  - familiar objects perceived as constant size (in spite of changes in visual angle when far away)
  - cues like overlapping help perception of size and depth
- What does this mean for items on the screen periphery?

Interpreting the signal (cont)

- 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
- Colour
  - made up of hue, intensity, saturation
  - cones sensitive to colour wavelengths
  - blue acuity is lowest
  - 8% males and 1% females colour blind

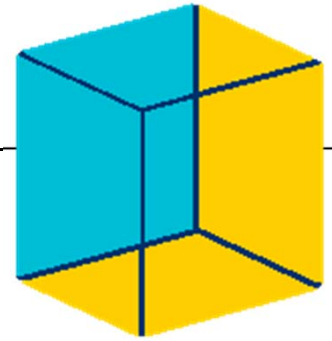
Brightness Contrast

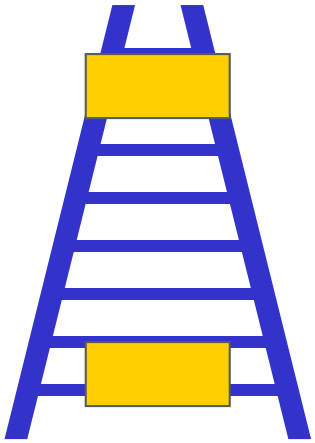


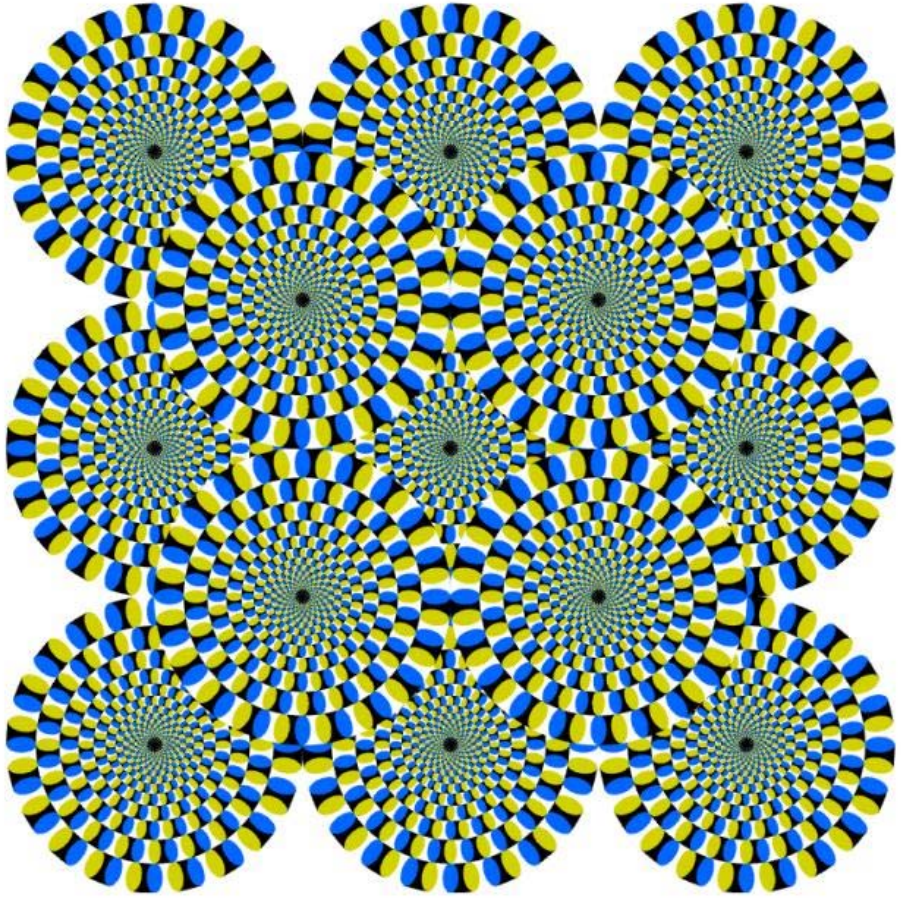
Interpreting the signal (cont)

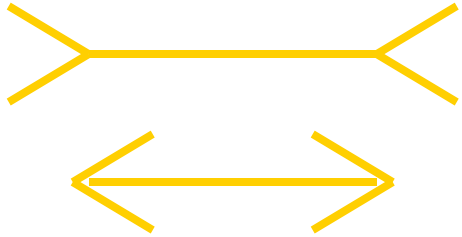
- The visual system compensates for:
  - movement
  - changes in luminance.
- Context is used to resolve ambiguity
- Optical illusions sometimes occur due to over compensation

Optical Illusions









http://www.ritsumei.ac.jp/~akitaoka/index-e.html

Visual Perception

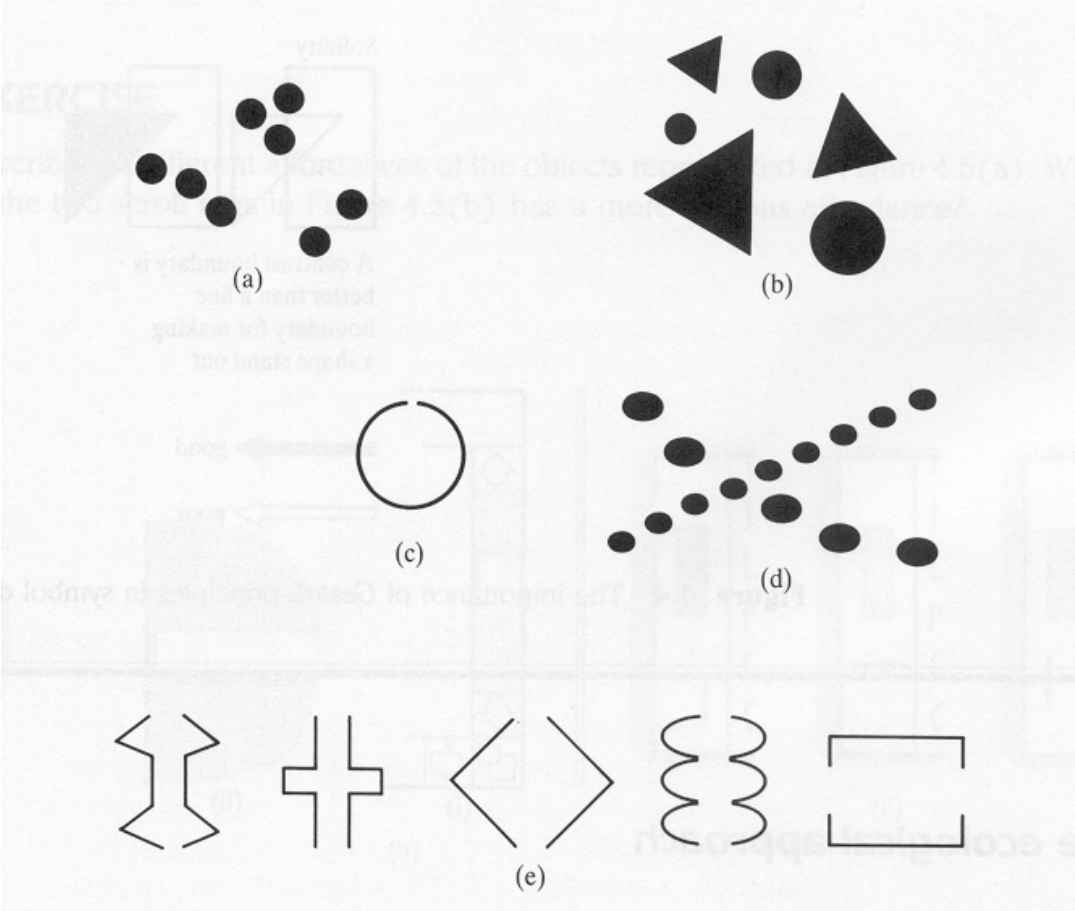
- Our most important sense
- Constructivist approach
  - Perception involves the intervention of representations and memories
  - Visual system constructs a model of the world by transforming, enhancing, distorting and discarding information

by R.C. James

Visual Perception

- Ecological approach
  - What do we do when we perceive?
  - What do we need to know about our environment to carry out our activities?
    - Users engage in activities that provide the necessary information
  - Affordances

Gestalt Laws of Perceptual Organisation



- Proximity
- Similarity
- Closure
- Continuity
- Symmetry

Concepts and Colors

8% of men and 1% of women are color blind

Green	%	Red	%	Yellow	%	Black	%	White	%
Safe	62.2	Hot	31.1	Caution	44.8	Off	53.5	Cold	71.5
Go	44.7	Danger	64.7						
On	22.3	Stop	48.5						

% of Hong Kong Chinese who associate particular concepts and colors (Courtney 86)

Green	%	Red	%	Yellow	%	Blue	%
Safe	61.4	Hot	94.5	Caution	81.1	Cold	96.1
Go	99.2	Danger	89.8			Off	31.5
		Stop	100				

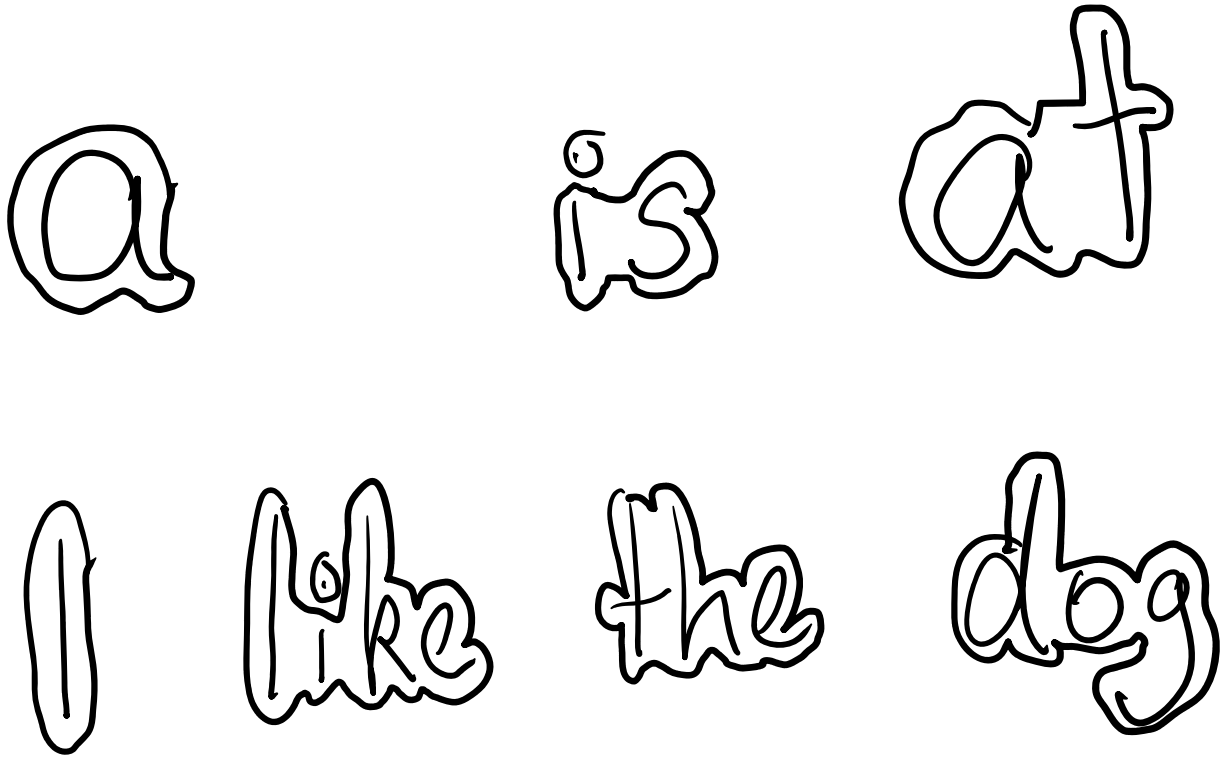
% of Americans who associate particular concepts and colors (Bergum&Bergum 81)

Reading

- Several stages:
  - visual pattern perceived
  - decoded using internal representation of language
  - interpreted using knowledge of syntax, semantics, pragmatics
- Reading involves saccades and fixations
- Perception occurs during fixations
- Word shape is important to recognition
- Negative contrast improves reading from computer screen

Word shapes

Always attend lectures



When you are surfing the web this week

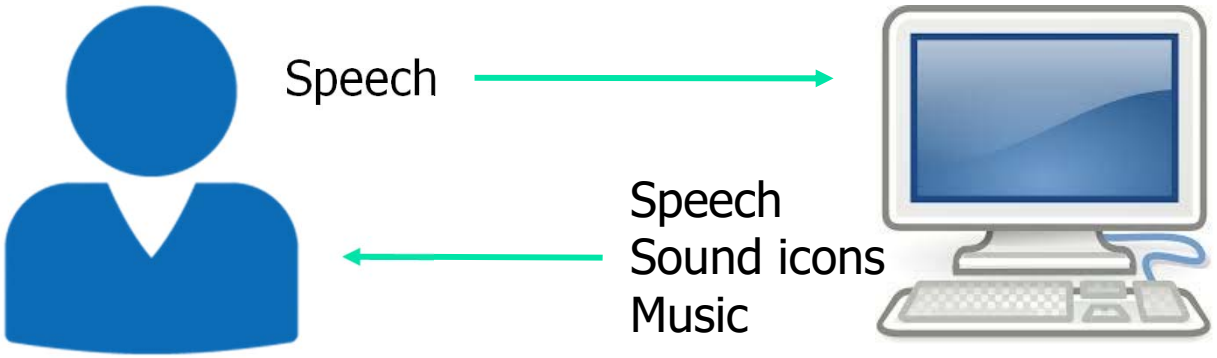
- If you have trouble finding something you know must be on a page

- Look to see why that is
  - Is it right on an edge?
  - Is the colour wrong?
  - Is the font too small?

What if ….

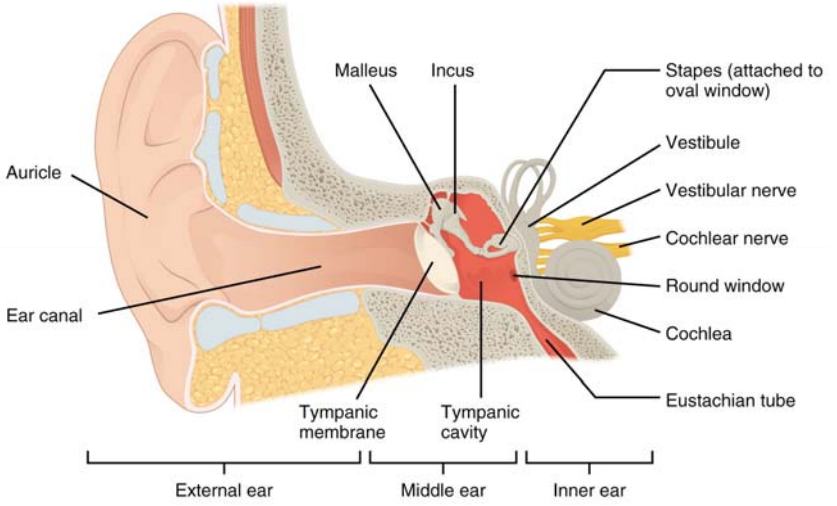
- Your visual attention is need for another activity
  - Driving and cell phone / GPS navigation….
- Are colour blind?
- Needed reading glasses?
- Had really poor eye sight that couldn't corrected by glasses?
- You are blind?

Hearing/Sound



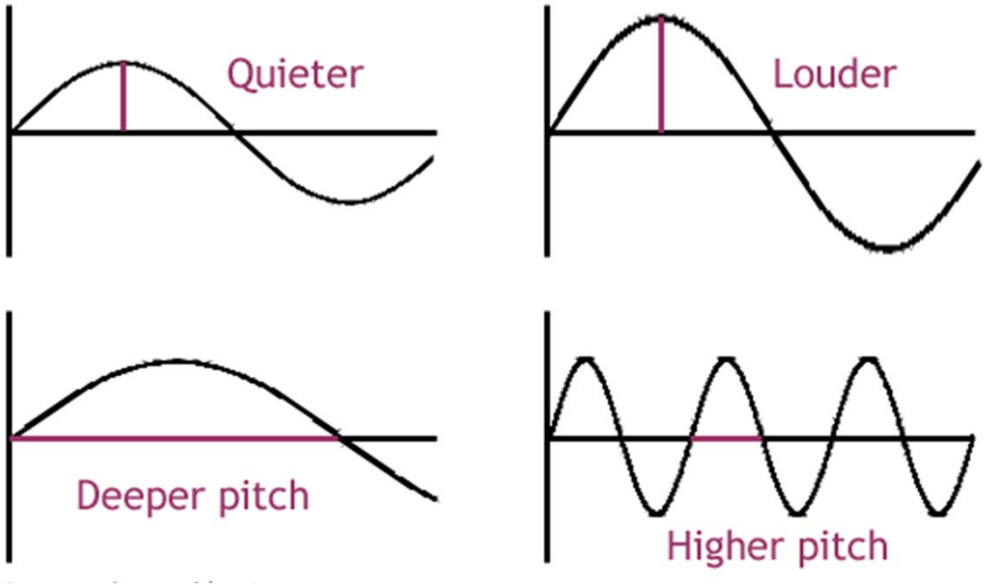
Human System Hearing

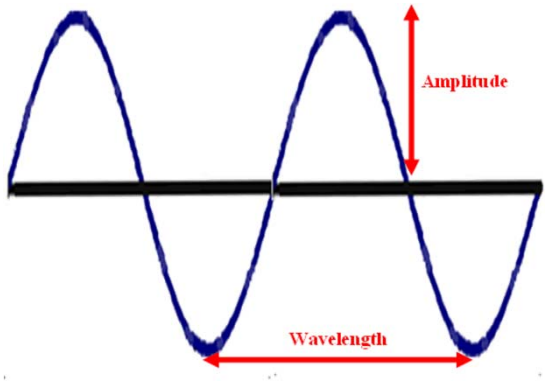
- Provides information about environment:
  distances, directions, objects etc.
- Physical apparatus:
  - outer ear – protects inner and amplifies sound
  - middle ear – transmits sound waves as vibrations to inner ear
  - inner ear – chemical transmitters are released and cause impulses in auditory nerve
- Sound
  - pitch – sound frequency
  - loudness – amplitude
  - timbre – type or quality
- Humans can hear frequencies from 20Hz to 15kHz
  - Less accurate distinguishing high frequencies than low
  - Higher frequency hearing disappears as you get older



https://cnx.org/contents/FPtK1zmh@8.25:fEI3C8Ot@10/Preface

Sound is vibration





Hearing (cont)

- Auditory system filters sounds
  - Can attend to sounds over background noise
    - The cocktail party phenomenon, more at: http://www.spring.org.uk/2009/03/the-cocktail-party-effect.php
  - Hearing aids disrupt this filtering
- Hearing is involuntary
  - Suddenly 'grabs' attention before we think
  - And some sounds are harder to ignore (e.g. baby crying)
- 'Listening' is voluntary (largely)
  - Whether we choose to process the meaning, especially if the sound is language (although something like hearing your name is pretty well involuntary)

What if....

- You are in a noisy environment
  - Night clubbing
    - Phone call/ text message?
- Your hearing is below average
- You are deaf

Computer Sound

- Computer Output/Generation (input to human)
  - Non speech
    - Music
    - Audio Icons (also called Earcons)
  - Speech

- Computer Input/Recognition
  - Speech
  - Non speech
    - Environmental
    - Music

Computer Output: Music

- Can be pre-recorded or generated
  - Movies
  - Games

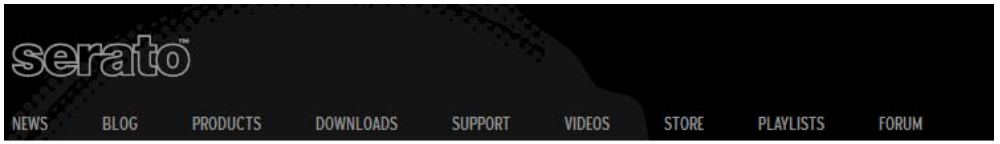
- Immersive experiences
  - Activates your brain in a different way from language
  - Acts almost entirely independently from hand-to-eye processing

Generating instruments

- MIDI – musical instrument digital interface
  - Allows very compact storage of music as the tones, durations and choice of synthesized instruments
  - Generally very 'computer generated' timbre/feel, see: http://www.midiworld.com/files/1128/
- Potential for much more sophisticated synthesized sounds for realistic or 'virtual' instruments
- Although the physics of real instruments can be quite complex
  - http://newt.phys.unsw.edu.au/jw/violintro.html
- Virtual: http://www.kurzweilai.net/instrument-of-the-future

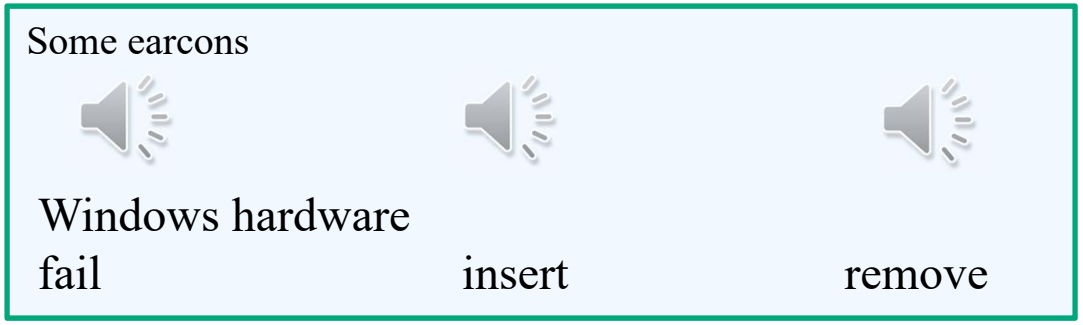
Generating music

- Exciting area for artists
- Everything from pseudo real to completely abstract
- There are Jazz music generators that only skilled people can differentiate from actual musicians.
- Serato – DJ software (www.serato.com)
  - Auckland company doing fantastic things
  - Several UoA grads there



Auditory Icons and Earcons

- Auditory icons: emphasis on 'natural' sounds and metaphor with real world; caricatures of naturally occurring sounds
  - e.g. sound of filling a bottle with water to match moving a large file
- Earcons: 'Artificial' sounds (generated)
  - e.g. more abstract metaphorical relationship to action or purely a convention (like corporate colour schemes)
- The difference between these two is subtle



Auditory Icons and Earcons

- Redundant Encoding
  - It aids memory by adding additional associations.
  - Can alert without interrupting (well, at least leaves the visual field clear)
  - An alterative communications channel.

- Positive/Negative Feedback
  - Auditory alarms might be crucial to the safe operation of computer-operated machinery or mission-critical environments
  - But too many alarms may be:
    - Annoying
    - Ignored
    - E.g., Three Mile Island.

Three Mile Island (1980s)

- Nuclear disaster in which there was a meltdown of the power station's nuclear reactor.
- Disaster could have been prevented with better designed control panels
  - A light indicated a valve had been closed when in fact it hadn't
  - The light indicator was obscured by a caution tag attached to another valve
  - The control room alarm system provided audible and visual indicators for more than 1500 alarm conditions. Evidently this number of alarms was intended to facilitate control of the entire plant during normal operating conditions. A single acknowledge button silenced all the alarms at the same time, but was not used because the operators knew they would lose information if they silenced some of the alarms

Using Sound in Interaction Design

- Learnability of the mapping between the icon and the object represented
  - "Oink" and "bow wow" have high articulatory directness (low distance between 'appearance' and function [or denotation])
  - A swishing sound accompanying a paintbrush tool also has high articulatory directness
  - A system beep, on the other hand, carries no information about what it denotes
    - But we may quickly learn to associate it with an error
    - And the square wave structure is a bit unpleasant, so it's better for an error than for feedback on success

Can you remember auditory icons?

- How many?
- How often do you hear them?

- Can you intuitively tell what these mean?

On

Off

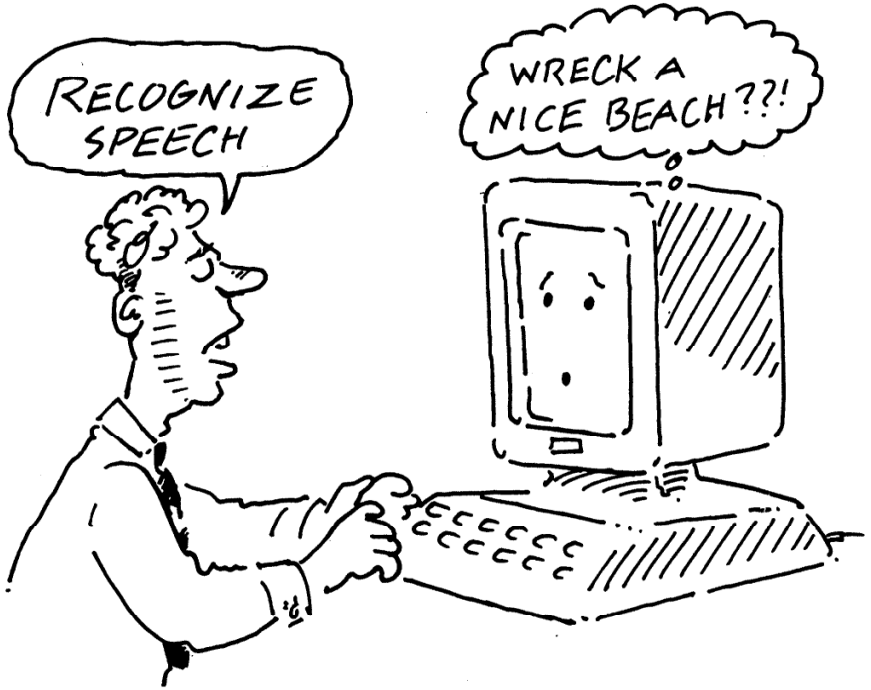
Sleep

Mis-recognized

Speech

- Where a person talks with a system that has a spoken language application, e.g. timetable, travel planner
- Used most for inquiring about very specific information, e.g. flight times or to perform a transaction, e.g. buy a ticket
- Also used by people with disabilities
  - e.g. speech recognition word processors, page scanners, web readers, home control systems

Have speech interfaces come of age?



Get me a human operator!

- Most popular use of speech interfaces currently is for call routing
- Caller-led speech where users state their needs in their own words
  - e.g. “I’m having problems with my voice mail”
- Idea is they are automatically forwarded to the appropriate service
- What is your experience of speech systems?

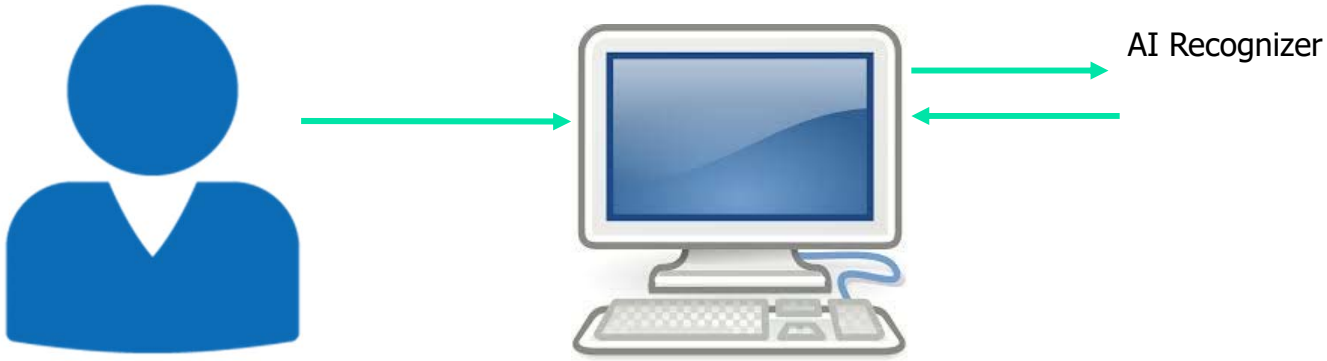
Format

- Directed dialogs are where the system is in control of the conversation
- Ask specific questions and require specific responses
- More flexible systems allow the user to take the initiative:
  - e.g. “I’ d like to go to Paris next Monday for two weeks.”
- More chance of error, since caller might assume that the system is like a human
- Guided prompts can help callers back on track
  - e.g. “Sorry I did not get all that. Did you say you wanted to fly next Monday?”

Research and design issues

- How to design systems that can keep conversation on track
  - help people navigate efficiently through a menu system
  - enable them to easily recover from errors
  - guide those who are vague or ambiguous in their requests for information or services
- Type of voice actor (e.g. male, female, neutral, or dialect)
  - Do people prefer to listen to, and will be more patient with, a female or male voice, a northern or southern accent?

How speech systems work



Dealing with errors with speech recognition

	Recognizer Correct	Recognizer Incorrect
Human Correct		
Human Incorrect		

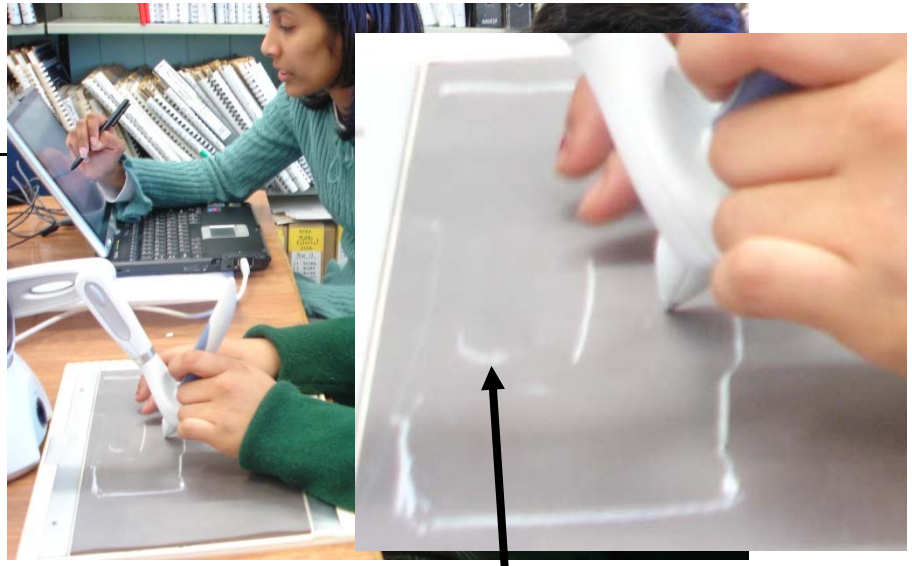
Touch

- Provides important feedback about environment.
- 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
    (some instant, some continuous)
- Some areas more sensitive than others e.g. fingers.
- Kinethesis - awareness of body position
  - affects comfort and performance.

Touch devices

- Interesting research in the areas of sound and touch
  - Prof Stephen Brewster

    http://mig.dcs.gla.ac.uk/



Pre-training circle

Post training circle

- The school has a PHANToM haptic device
  - This is what is being used for the 'teaching visually impaired kids to sign their name' project
  - https://www.cs.auckland.ac.nz/~beryl/publications/

Movement

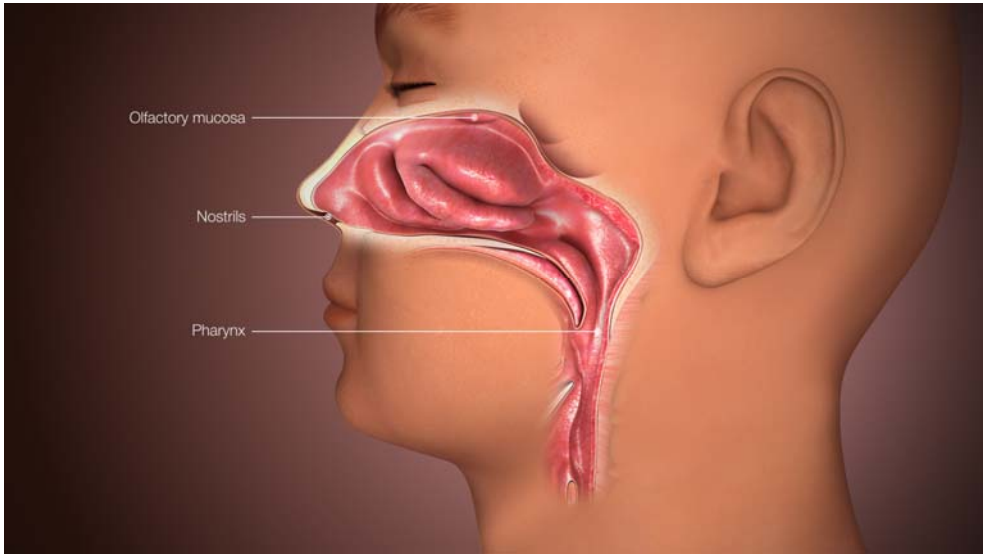
- Time taken to respond to stimulus:
  reaction time + movement time
- Movement time dependent on age, fitness etc.
- Reaction time - dependent on stimulus type:
  - visual ~ 200ms
  - auditory ~ 150 ms
  - pain ~ 700ms
- Increasing reaction time decreases accuracy in the unskilled operator but not in the skilled operator.

What if…

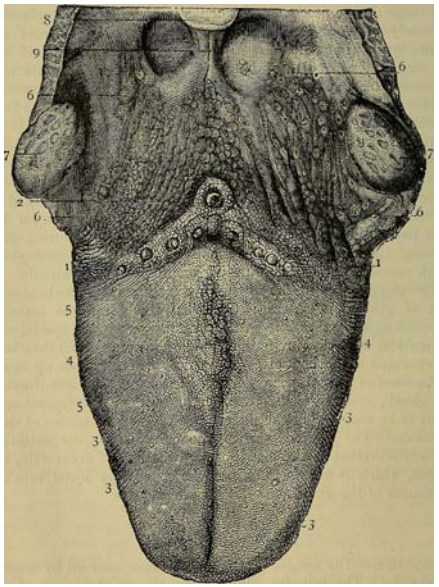
- You can't keep your hand steady?
  - Keyboard, mouse
- You are a paraplegic?
- http://www.abilityhub.com/mouse/eyegaze.htm
- Eye tracking software/hardware also use for usability studies to track users focus points

Smell and Taste

- We have about 4000 types of different smell receptors
- Some primitive attempts to analyse and synthesise smell
- It is technically very difficult!
- Taste is closely associated – little work in this area



By https://www.scientificanimations.com
https://www.scientificanimations.com/wiki-images/, CC BY-SA 4.0



By Internet Archive Book Images –
https://www.flickr.com/photos/internetarchivebookimages/14578806517/

Summary

- The *Model Human Processor* gives insight into human abilities
- Extraordinary though our senses are, there are constraints on our capabilities
  - Human memory limitations should be considered
- Understanding the capabilities of our senses informs UX design

Multi-choice

You decide to enhance your tablet-based sketching system by using redundant encoding for some functions. To match the visual feedback picking up a new pen to draw with the best approach would be to use:

a) Auditory icon
b) Vibration
c) Earcon
d) Flashing icon in the user's peripheral vision