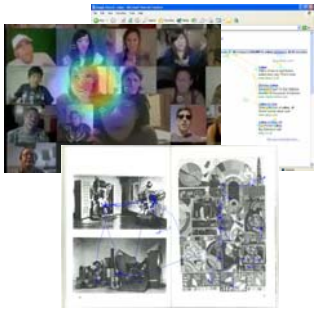
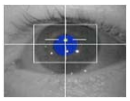
Case Studies in Design Informatics
Lecture 12
Affective Computing: Gaze



Robin Hill
Institute for Language, Cognition & Computation, School of Informatics



Lecture 12 content

Visual environment

Attention **Behaviour** **Response**

Input recognition (Lecture 9)

Output synthesis (Lecture 10)

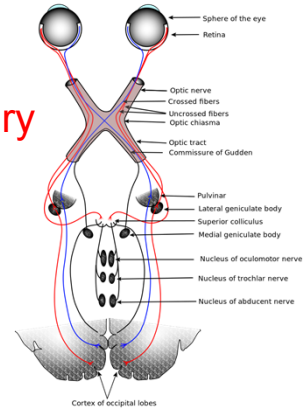
Main part

Interocular Trauma Test (IOTT)

Some things just hit you between the eyes when you see them.

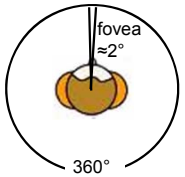
Visual system

Brain hungry
&
complex



Human vision is very focused

We only see a small portion of the world in detail at any one moment!



Simple stimuli: text (lect. 11)

- Fovea: central 2° of vision
- Parafovea: from foveal region to about 5.0º from fixation
- Periphery: everything beyond parafoveal region

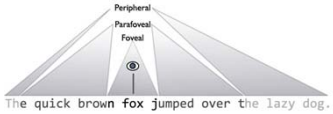


Fig. 1 The foveal, parafoveal, and peripheral regions when three characters make up 1° of visual angle. The eye icon and dotted line represent the location of fixation

From Schotter et al. (2012)

Implication

1. Large proportion of the brain dedicated to visual processing.

\+ 2. Limited momentary input.

= 3. Where you are looking *now* is very important and highly informative.

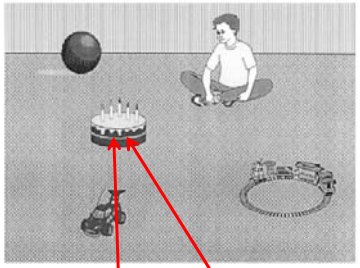
Eye movements

- Eyes have the most active muscles in the body.
- Movements even when asleep (REM sleep).
- Fixations (stable, typically about 3 or 4 per second).
- Saccades (ballistic jumps – visual suppression).
- Smooth pursuit (tracking moving objects).
- Pupillometry (pupil dilation as an indication of affection: emotional arousal and cognitive load – not just a reaction to light).

Cognitive processing

- Eye-movement data are a good moment-to-moment indicator of visual-cognitive processing
- Strong Eye-Mind Hypothesis
  - Just & Carpenter (1980)
- *Where?* (spatial)
  - Eye movements as indicators of overt visual attention: **where** we currently **look at** (e.g., word in a sentence, object in a scene) is **where** we currently **attend to**
- *When?* (temporal)
  - Fixation durations: increased processing difficulties show in longer fixation times

Anticipation in language

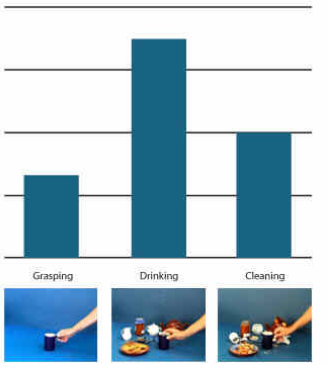


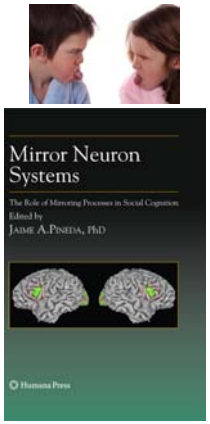
While viewing the scene, participants heard:
1) The boy will move the cake; or
2) The boy will eat the cake

(Altmann & Kamide, 1999)

Gesture

Mirror neurons



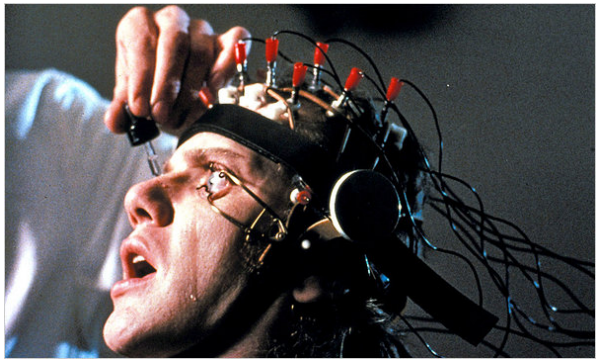


Hardwired empathy

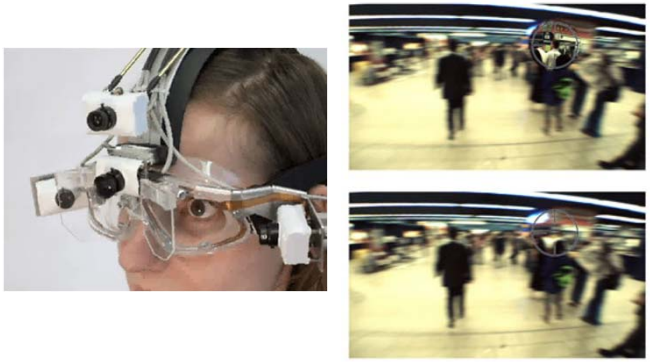
- Observing the same action in different contexts elicits different levels of mirror neuron activity.
  - Suggests that the mirror neuron system does more than code the observed action ("that's a grasp"). It also codes the intention behind the action ("that's a grasp to drink" or "that's a grasp to clear the table"). [UCLA School of Medicine.]
- Instrumental for interpreting the facial expressions and actions of others.

Gaze data

Eye-tracking technology



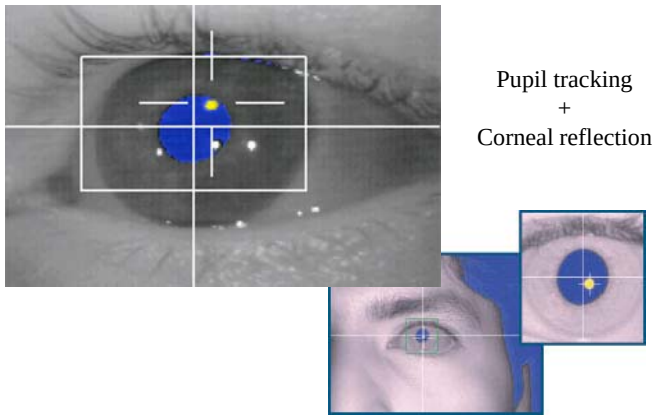
Eye tracking in the real world



"EyeSeeCam" (LMU Munich; Schneider *et al.*, 2009)

Eyelink trackers

Infra-red video monitoring



Pupil tracking
+
Corneal reflection

Eyetracker (geek) output

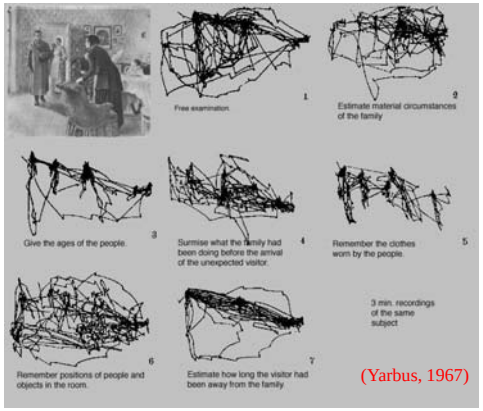
```
        x     y          x     y
825554 77.7 219.7 1037.0 79.2 249.5 1088.0
825555 77.1 219.1 1037.0 79.3 249.1 1088.0
825556 77.0 218.6 1037.0 79.1 247.9 1087.0
825557 77.0 218.6 1037.0 79.2 247.0 1086.0
825558 77.2 218.6 1037.0 79.3 246.2 1085.0
825559 77.4 218.6 1036.0 78.8 245.8 1086.0
825560 77.4 219.0 1036.0 78.1 245.4 1086.0
825561 77.5 219.4 1036.0 77.4 244.9 1087.0
825562 77.5 219.9 1036.0 77.6 244.8 1086.0
825563 77.4 219.8 1036.0 77.9 245.3 1085.0
825564 77.1 220.0 1036.0 78.2 246.3 1085.0
825565 77.0 220.3 1036.0 78.3 247.3 1086.0
825566 77.4 220.5 1037.0 79.0 247.5 1088.0
825567 77.7 220.5 1037.0 80.1 247.2 1089.0
825568 78.3 220.5 1037.0 81.2 246.8 1088.0
825569 79.0 220.5 1037.0 82.4 245.3 1087.0
825570 79.8 220.5 1037.0 83.3 243.9 1087.0
825571 80.6 220.5 1037.0 84.5 243.4 1087.0
```

- column 1: Time stamp (msec, internal eye-tracker clock)
- columns 2-4 = **Left** eye:
- column 2: *Horizontal* gaze position (pixels)
- column 3: *Vertical* gaze position
- column 4: Pupil diameter or area value
- columns 5-7 = **Right** eye:
- column 5: *Horizontal* gaze position
- column 6: *Vertical* gaze position
- column 7: Pupil diameter or area value

The Unexpected Visitor
Ilya Repin (1884) [artist]

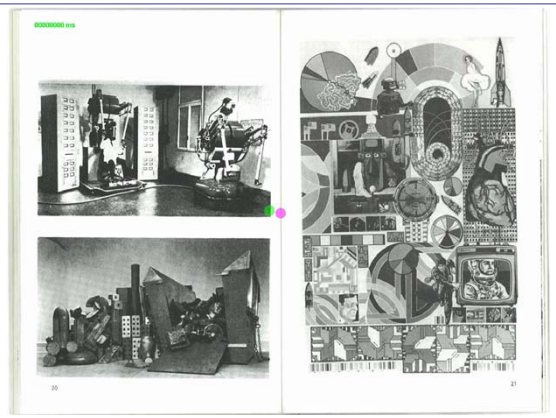
(Yarbus, 1967)

Task effects



(Yarbus, 1967)

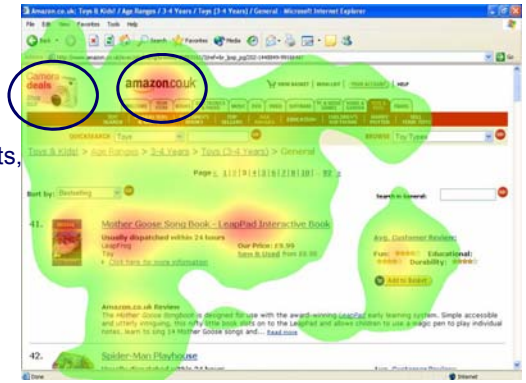
Two eyes on Paolozzi



Browsing / interface design

Browsing / interface design

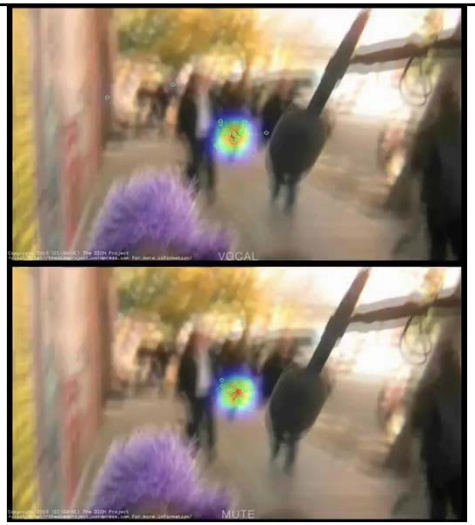
Adverts, logos, etc.



Browsing / interface design

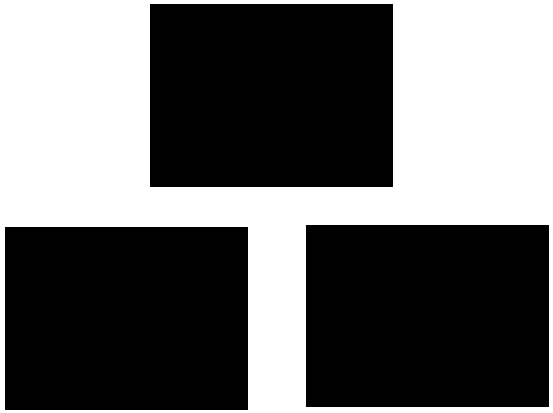
Interaction: competitive play

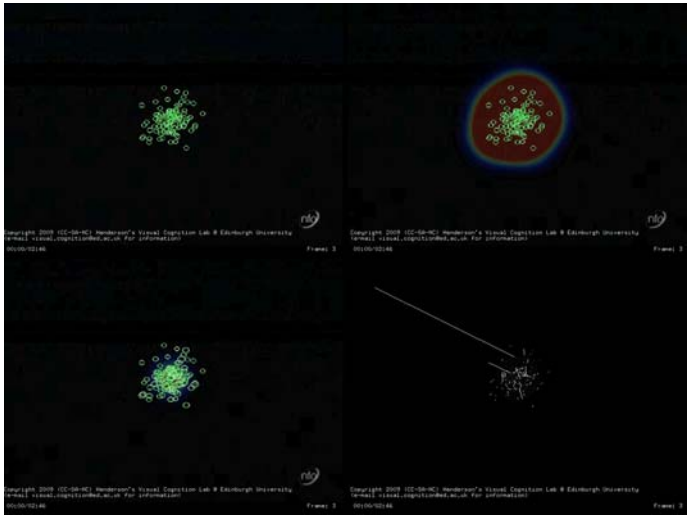
Social interaction



Face-to-face communication

https://vimeo.com/visualcognition





Pupillometry

- Changes in pupil dilation due to psychological processes.
- Dilation is not just a reflex to light.
- Reaction to arousal or emotions (positive or negative).
- Evidence of task demand (cognitive load).

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, 59(1-2), 185-198.
Granholm, Eric, & Steinhauer, Stuart R. (2004). Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology*, 52(1), 1-6.
Scheepers, Christoph, Mohr, Sibylle, Fischer, Martin H., & Roberts, Andrew M. (2013). Listening to limericks: A pupillometry investigation of perceivers' expectancy. *PLoS ONE*, 8(9), e74986.

Monitored behaviour and dialogue

Neuromarketing

http://www.smivision.com

Physiognomy

- System for identifying personality types (incl. criminals) from facial characteristics.
- Aristotle: “noses with thick, bulbous ends belong to persons who are insensitive, swinish”.
- Too extreme! Debunked by science.

Illustration from a 19th-century book on physiognomy. [Encyclopaedia Britannica]

Automatic facial expression analysis software

Can automatically detect:

- Valence (positive, negative or neutral responses).
- Emotions (joy, anger, surprise, fear, sadness, disgust, contempt, frustration and confusion).
- Facial landmarks (e.g. corners of mouth)
- Various set actions/behaviours.
- Head pose.

Goldberg, Joseph H. (2012). Relating perceived web page complexity to emotional valence and eye movement metrics. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1), 501-505.

The Emotient Module

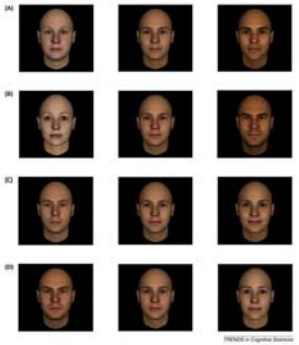
Image showing the integration of stimuli, face of the participant and Emotient emotion channels in time line.

http://imotionsglobal.com

Hi-tech full monty

https://vimeo.com/64852896

Synthesis

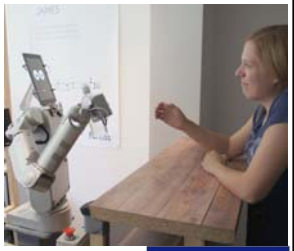


Faces generated by data-driven computational models of evaluations of **(A)** competence, **(B)** dominance, **(C)** extroversion, and **(D)** trustworthiness. The face in the middle column represents an average face in the statistical model. Faces in the right column are 3 standard deviations (SD) above the average face on the respective trait dimension; faces in the left column are 3 SD below the average face on that same dimension.

Olivola, Christopher Y., Funk, Friederike, & Todorov, Alexander (2014). Social attributions from faces bias human choices. Trends in Cognitive Sciences, 18(11), 566-570.

Example: Joint Action

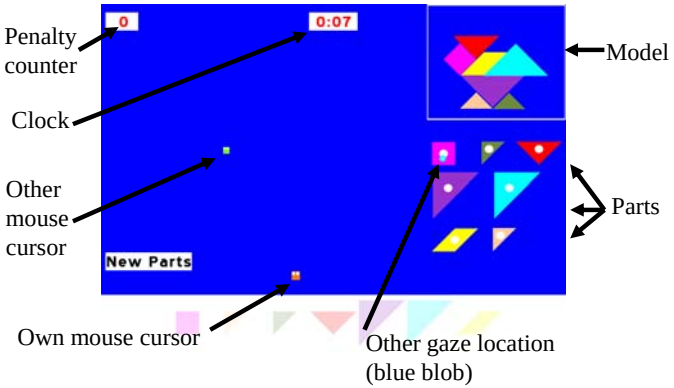
Multimodal: combines eye movements, motor control (actions), problem solving and dialogue.



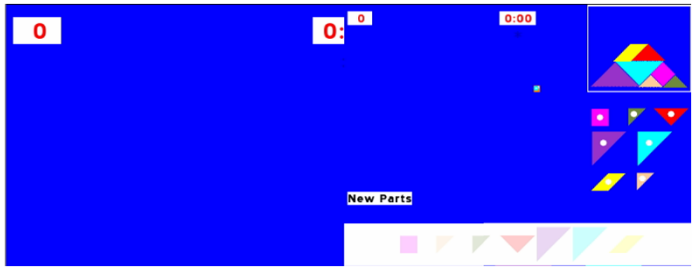
JAST Joint-Action Science and Technology

http://wcms.inf.ed.ac.uk/jast

Human collaboration



Parallel video example



Richness of human language

"I'll hold on to this bit and you put the hat on."

1. The listener is not about to place something on their head.
2. There is no hat. This refers to a red triangle that is at the top of the model.

- The person's partner had no problem understanding this but how easy is it for an artificial agent?
- Visual context!

CASMACAT

Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation

Multinational, multidisciplinary EU Project

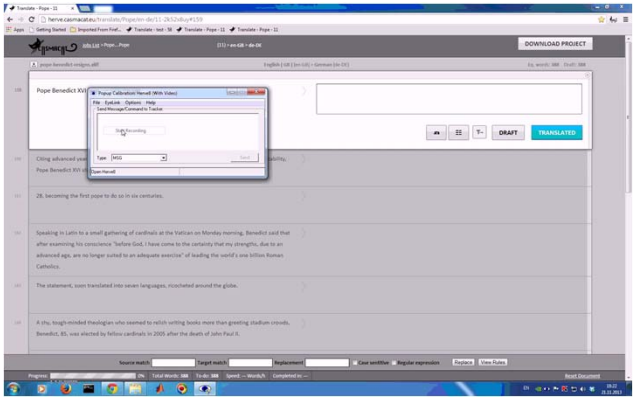
www.casmacat.eu

Reading comprehension
Writing production
Bilingualism
User Interface (UI) design
Expert behaviour

Extreme examples but it is always worth having a human with knowledge of the target language check any Machine Translation output.

- Affective – but not as intended.
- Mistakes are costly and embarrassing.

Human-Computer Interaction



Anomaly detection

No error version (total reading times on each word in ms)

Steve	Jobs	was	certainly	a	designer,	a	dreamer	and	self-proclaimed	Hippie.
231	453	487	1254	151	624	0	1097	192	2624	688

Steve Jobs was certainly a designer, a dreamer and self-proclaimed Hippie.

Error version (total reading times on each word in ms)

Steve	Jobs	were	certainly	a	designer,	a	dreamer	and	self-proclaimed	Hippie.
434	590	1862	538	0	403	0	421	0	810	302

Steve Jobs were certainly a designer, a dreamer and self-proclaimed Hippie.

Expertise (user groups)

- Compared a group of professional translators and a group of non-professionals (fluent bilinguals).
- Larger increase in pupil dilation for the non-professional group when detecting errors in translation.
- The same task required more intense effort and engagement for novices, even if they achieved the same final result.
- Pupillometric evidence that experts found the task easier and were more relaxed.

Finished! Happy or sad?