# Chapter 3:

## Humans: Stereo Vision, Reading, Hearing, Space, Territory and Emotions

## Overview

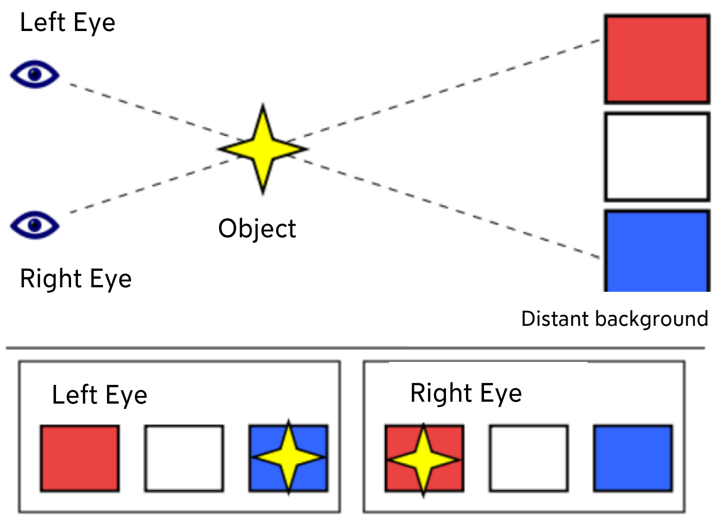


## 1. Stereo Vision

Everything on a 2D display is 2D! If we see it three dimensional, we imagine it. For example, when displaying a projection of a 3D model on a 2D screen. We can interpret this as 3D because we have experience with this.

“Real” 3D, however, requires an image for each eye. This happens “naturally” when looking at 3D objects in physical space. But it can also be simulated by providing a separate image for each eye using technologies that can provide 3D content.

The basis for this technology is the so-called **parallax**. It describes a displacement or difference in the apparent position of an object viewed along two different lines of sight and is measured by the angle or semi-angle of inclination between those two lines.

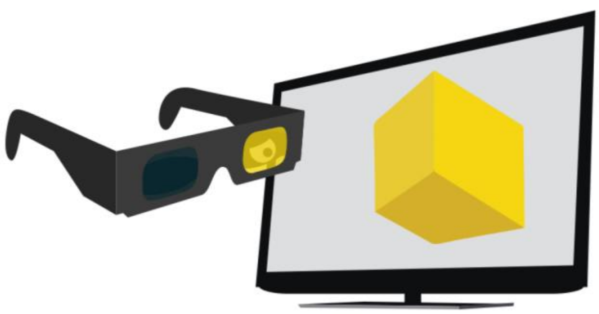



To display these two different images for the eyes we want to present three different methods:

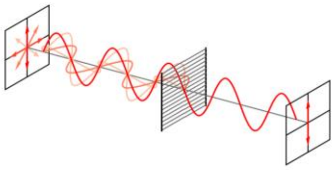
- **Shutter systems**
- **Polarized systems**
- **Virtual reality headsets**

<u>Shutter systems</u>

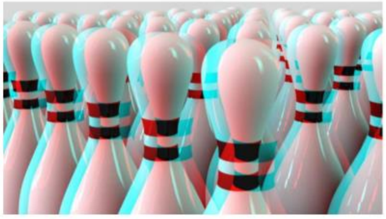
Shutter systems consist of glasses that are synchronized with a monitor. The glasses block alternately the left and the right eye. Synchronous to that, the monitor alternately displays the image for the left and right eye. This switching between left and right eye happens at a very high frequency so that the image is perceived as continuous. In that way, two different images can be delivered.



Polarized systems



Polarized systems are the most popular systems, since modern cinemas or 3D Monitors use this technology, too. The image for the left and right eye is decoded in only one visual wave direction/angle. The polarized glasses have filters that let the correct wave direction pass the filter. In that way, only the appropriate image is sent to the corresponding eye. This technology is cheapest; however, it has limitations. If the user does not look from the correct viewing angle and moves the head too much, the illusion of 3D fails.



Virtual reality headsets

Virtual Reality Headsets have two distinct displays for each eye. In that way, you can deliver two different images at very high frequencies and quite high resolution. Modern VR-headsets have resolutions up to 1832x1920 pixels per eye and even higher rendered at 90 Hz and higher. This leads to a huge workload. That is why some Headsets are still connected to a powerful PC with a good graphics card.

## 2. Reading

Reading consists of several stages:

- Visual pattern perceived
- Decoded using internal representation of language
- Interpreted using knowledge of syntax, semantics, pragmatics

Reading also involves saccades and fixations. Thus, our eyes fixate a word and when the eyes move to the next word, we perform a saccade. The perception of words occurs during fixations. For a proper recognition the word shape plays an important role. Sometimes words are capitalized to emphasize them. However, it was found that all capitalized words are harder to read and thus decrease your reading spead.

https://uxmovement.com/content/all-caps-hard-for-users-to-read/
http://en.wikipedia.org/wiki/Reading_%28activity%29

Some basic facts about reading:

- Typical reading speeds are 100 (memorizing) to 1000 (scanning) words per minute
- Reading skills differ to a great extend (according to PISA more than 20% have difficulties in reading)
- Reading speed has for many tasks a significant impact on overall user performance
- Good readers "recognize" words (they do not read them letter by letter)
- Providing a visual presentation that supports reading is important (font, size, color, length of lines, structure, ...)
- Reading from a computer screen is in general slower than from paper

Interestingly the order of the letters in a word is not too important to read and understand a text. Read through the following text and try to understand it:

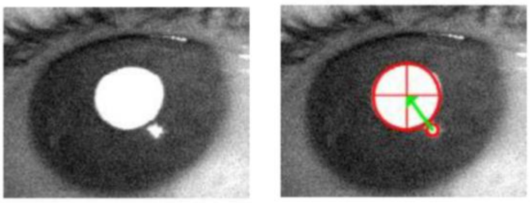
> I cnlduo't bvleiee taht I culod aulaclty uesdtannrd waht I was rdnaieg. Unisg the icndeblire pweor of the hmuan mnid, aocdcrnig to rseecrah at Cmabrigde Uinervtisy, it dseno't mttaer in waht oderr the lterets in a wrod are, the olny irpoamtnt tihng is taht the frsit and lsat ltteer be in the rhgit pclae. The rset can be a taotl mses and you can sitll raed it whoutit a pboerlm. Tihs is bucseae the huamn mnid deos not raed ervey ltteer by istlef, but the wrod as a wlohe. Aaznmig, huh? Yaeh and I awlyas tghhuot slelinpg was ipmorantt! See if yuor fdreins can raed tihs too.

**What is the only rule for the letters in words so that we can still read them? How could this influence UI design?**

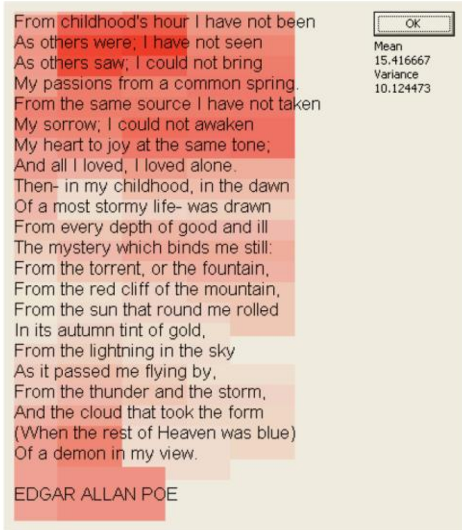
The basic facts and such phenomena are found by conducting studies and observing how people react. But especially when we want to know how people read, we would have to know where the person is looking at. This can be done using **eye-tracking**. Eye-Tracking allows to see where someone is looking at while reading a text.
The principle of a video-based eye gaze tracking system is shown here:

- The picture below shows the camera with an infrared LED mounted below the tablet PC.

- The white pupil in the camera image comes from reflection of infrared light
- The infrared light also causes a reflection glint, which does not move as the eye moves
- The position of the gaze on the screen can be calculated by the distance from the glint to the pupil center



When using eye-tracking for reading analysis, you can generate the following image overlays, that show where the person is/was looking at while reading the text:


You can also generate heat maps that color code where people were looking at the most:

From these images you can draw new conclusions like:

- Users first read in a horizontal movement
- Users move down the page a bit and then read across in a second horizontal movement
- Finally, users scan the content's left side in a vertical movement.

Eye-tracking is of course not limited to reading analysis. Potential other application areas include:

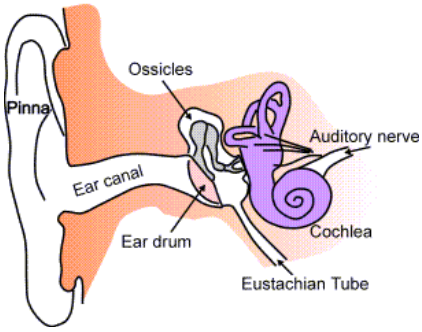
- Clinical Research
- Marketing and Consumer Research
- Infant and Child Research
- Education
- Human Performance

# 3. Hearing, Touch, Movement

**Hearing**

Some basic facts about our sense of hearing:

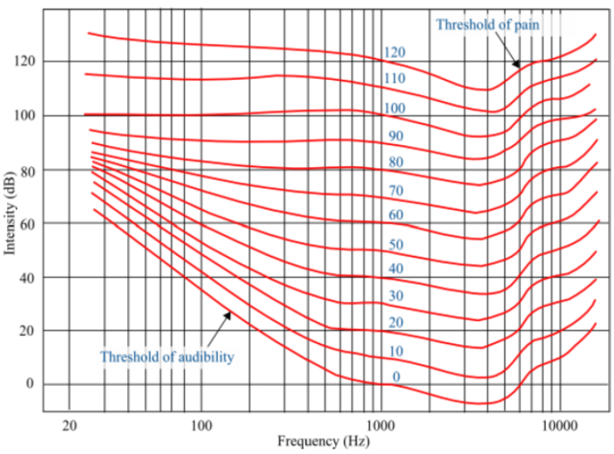
- We have two ears with which we collect information about environment, evaluate the type of sound source, evaluate the distance and direction
- The physical apparatus (ear) consists of:
    - Outer ear – protects inner ear, amplifies sound (3-12 kHz)
    - Middle ear – transmits sound waves as vibrations to inner ear
    - Inner ear – chemical transmitters are released and cause impulses in auditory nerve
- The sound that we can capture with our ears consist of:
    - Pitch – sound frequency
    - Loudness – amplitude
    - Timbre – type or quality




From experience you could know that very high pitches are oftentimes uncomfortable and lead to pain. You automatically cover your ears with your hands to lower the loudness of the sound. This is also described by the threshold of pain. So, a certain loudness level above which the sound leads to pain. There is also a threshold which must be passed to hear a sound at all, the threshold of audibility. These two thresholds can also be seen in the **Fletcher-Munson equal-loudness contours**:

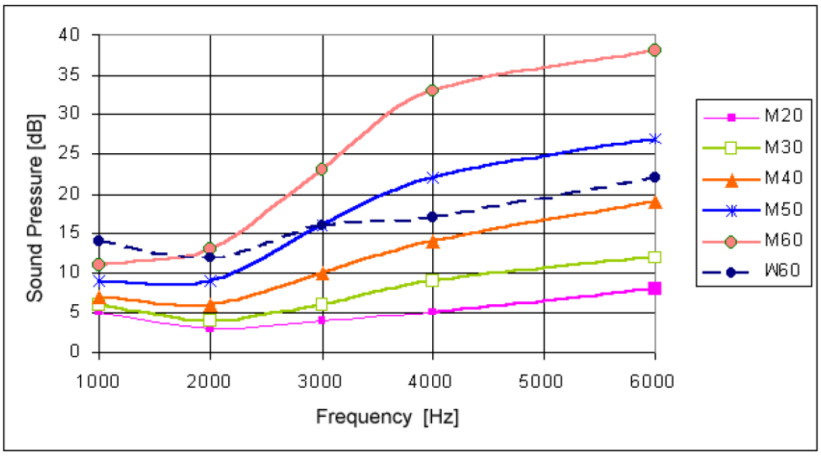
+ http://en.wikipedia.org/wiki/Absolute_threshold_of_hearing

These curves describe the perceived loudness of a generated sound. For example, the red curve in the middle shows a "60" at about 1000 Hz. So, it is a sin wave with intensity of 60 dB. When increasing or decreasing the frequency of that sine wave the perceived loudness of that sound changes. The loudness that would be required to have an equal impression of loudness can be read from the red curves. For example, for a low frequency of 30 Hz you need approximately 80 dB for the same 60 dB loudness impression. In that way it was possible to understand how people hear and define the thresholds of hearing and pain.




https://blog.landr.com/fletcher-munson-curves/#:~:text=What%20are%20Fletcher%2DMunson%20curves,its%20frequency%20for%20human%20liste

Additional to that, our sense of hearing changes with increasing age and becomes worse. Especially the high frequency perception is affected. The following graph depicts this process. The older you get, the higher sound pressure is required to perceive high frequencies at all. The low frequency perception is not affected that severe.




**Why is it important to know about the human sense of hearing and how it changes with frequency and age?**
**What are the implications for user interface design?**

Another interesting phenomenon of our sense of hearing is the **selective hearing** often related to as *cocktail party effect*:

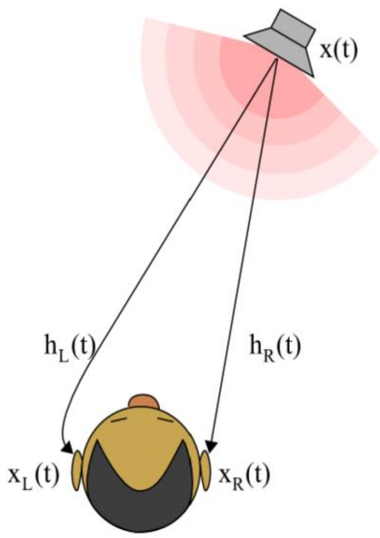
- You are in a noisy environment like a crowded underground train, and you can still have a conversation. You can even direct your attention to another conversation and "listen in".
- You are in a conversation and somewhere else someone mentions your name. You realize this even if you have not been listening actively to this conversation.

The auditory system filters incoming information and allows selective hearing of sounds in environment with background noise or keywords. This is a **binaural** effect. Thus, we need both ears for that to identify the sound source's location.

To locate the sound source, we can rely on three effects:

- Interaural time difference (ITD)
- Interaural intensity difference (IID)
- Head related transfer functions (HRTF)

ITD describes the time difference of the sound on arrival at the different ears. IID describes the sound pressure difference on arrival at the different ears. HRTF describes how the head changes the sound because of masking.



While ITD and IID can be calculated quite easily, HRTFs are more complex and must be measured in expensive experiments. The reason for measuring HRTFs is to create a better experience for 360° sound or stereo signals when listening to music or for other applications like in Virtual Reality. If you simply play the stereo sound you generated in earphones there is no body/head that changes the sound perception (masking, damping, …). Hence, it is pre-calculated in the signal played. This requires measurements with a dummy head (e.g. microphone at the position where typically the ear is). Based on the data a function can be developed.

**Touch and Movement**

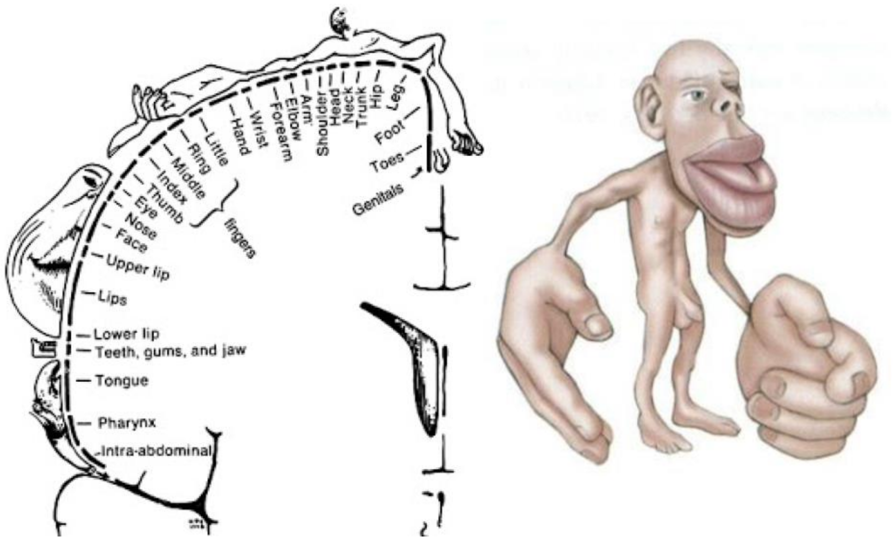
Our sense for touch and movement provides important feedback about our environment. It is said to be the key sense for someone visually impaired. The environmental stimulus is received via receptors in the skin, which are:

- Thermoreceptors – heat and cold
- Nociceptors – pain
- Mechanoreceptors – pressure (instant / continuous)

Some areas more sensitive than others (e.g. fingers compared to our back) because the density of those receptors is higher in fingers. Based on the "raw" data we collect through those receptors (and also other receptors inside muscles and the like), we are able to further process the information in our brain and get a feeling of our body in space. Thereby we can define two holistic body perceptions:

- Kinesthesis: feeling of limb and body movements
- Proprioception: unconscious perception of movement and spatial orientation arising from stimuli within the body itself.

As we already mentioned the fingers are more sensitive than our back. This is also represented in the so-called **Somatosensory Homunculus**. It is shown where the body parts are mapped onto the surface of the brain. The larger the brain region the more sensitive the body region. Based on this mapping on the brain's surface the homunculus was created to show the imbalance between real body size and sensitivity.




http://tmww.blogspot.com/2011/05/homunculus-of-touch.html

## 4. Space and territory

Humans use space to ease tasks (simplify choices, perceptions, and internal computation). However, computer systems often do not support this well.

> *'How we manage the spatial arrangement of items around us is not an afterthought: it is an integral part of the way we think, plan, and behave.'*
> David Kirsh. The Intelligent Use of Space. Artificial Intelligence (73) Elsevier, p31-68, 1995

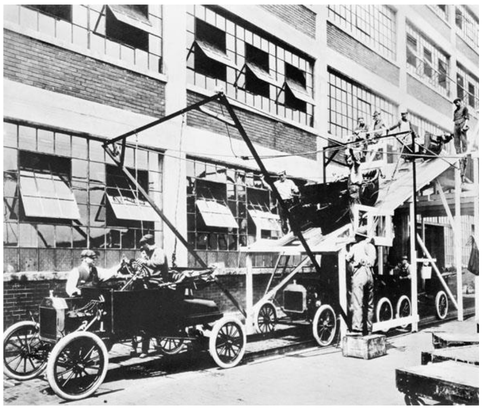
When space is used efficiently than some effects are:

- Reduced cognitive load (space complexity)
- Reduced number of steps required (time complexity)
- Reduced probability of errors (unreliability)

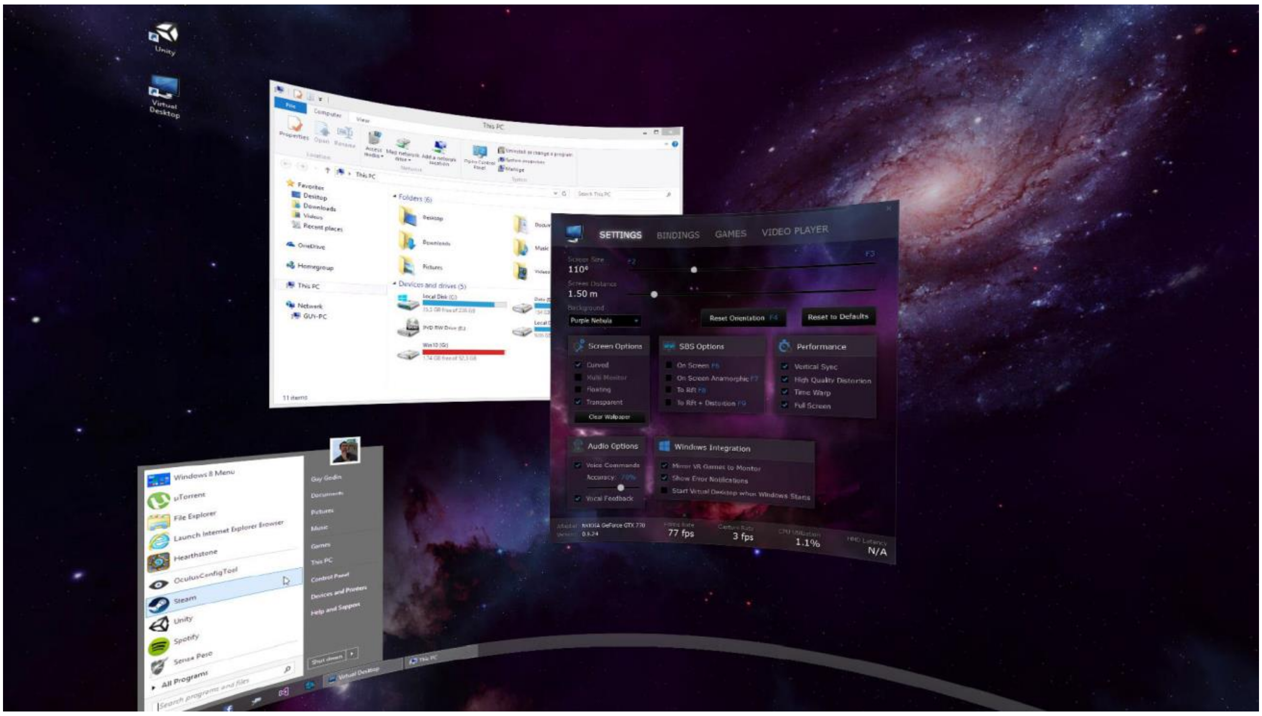
There are some general rules for the intelligent use of space:

- Utilize space as much as possible
- Use space in physical world / on screen
- Allow users to customize spatial arrangements
- Provide interactive means for manipulation of objects in space
- The physical space + spatial order:
    - Implies behavior
    - Eases categorization
    - Allows to make (internal human) computation easier
- Segment problems and tasks
    - Spatially
    - Temporally

Especially the last two bullets (The physical space + spatial order as well as segment problems and tasks) can be seen with the invention of Ford's first assembly line. One worker always has to do one thing. So the place in the line implies the behavior of that worker. Perhaps at the end of the line he/she simply has to mount the door handles and the car is finished. By dividing these task across the assembly line it was possible to speed up the process by implied behavior for a worker, eased categorization (crew for machine, crew for chassis, …), eased computation because crews only have one specific task and don't have to think of the complete and more complex car. The assembly line also devides the task automatically in spatially and temporally different zones/phases.



**What do you think about the following interface and how is it handling space?**




In computer science or software in general there are some UIs that are equivalent to an assembly line. Think about the following:

- Wizards
- Guided tours
- (Distributed) workflow
- Tools that have support for different roles
- User interfaces that restrict choice as appropriate for a given context
- Different applications for different tasks
- Different work environments for different tasks (e.g. CAD workstation, video editing station, POS terminal)

## 5. Emotions

There are various theories of how emotion works:

- James-Lange: emotion is our interpretation of a physiological response to a stimuli “we are sad because we cry…”
- Cannon: emotion is a psychological response to a stimuli
- Schachter-Singer: emotion is the result of our evaluation of our physiological responses, in the light of the whole situation we are in

Despite the various theories, one can say that emotion clearly involves both cognitive and physical responses to stimuli.

The biological response to physical stimuli is called **affect**. The Affect influences how we respond to situations:

- Positive ➔ creative problem solving
- Negative ➔ narrows thinking

“Negative affect can make it harder to do even easy tasks; positive affect can make it easier to do difficult tasks” (Donald Norman).

**Why is emotion important? What are 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

Especially the last point was tested with ATM machines. The experiment consisted of six ATM identical in function and operation. Some were aesthetically more attractive than others. The result of the experiment was that the nicer ATMs are easier to use… So, aesthetics can change the emotional state and emotions allow us to quickly assess situations. Positive emotions make us more creative and attractive things make feel people good. (D. Norman, Emotional Design, Chapter 1)

Another theory is called the **Affordance Theory**. It describes the (perceived) possibility for action. The original idea was stated by Gibson. He said that objective properties imply action possibilities. So, how we can use things – independent of the individual person. Norman later added the perceived affordance also includes experience of an individual.

Gibson, J.J. (1979). The Ecological Approach to Visual Perception, Houghton Mifflin, Boston.
Norman, D. A. (1988). The Psychology of Everyday Things. New York: Basic Books. (The paperback version is Norman, 1990.)

A simple example is vandalism at a bus stop…
Is the bus stop built out of **concrete**, you will surely find graffiti on it.
Is the bus stop built out of **glass**, it will be smashed.
Is the bus stop built out of **wood**, you'll find carvings.

These examples just give the impression that affordance is not only for HCI but can be found everywhere. The implication for affordance is to build natural and intuitive user interfaces that already by design imply how they must be used.

A good example is a full body interaction mimicking a real action like tennis or swinging a sword or moving your hands. These interactions can also be used by elderly people that oftentimes are not used to working with PCs.