HCI Assignment

Srikumar Sastry (201601016)

Q 1) What is the attention? Define top-down and bottom-up attention with examples. What is multi- tasking in relation to attention? Elaborate the implication of attention in human-computer interface designing.

Attention, also known as focussed or directed consciousness, is a complicated procedure carried out via recreation in various neurophysiological systems. Attention has advanced to allow the character to correctly cope with his environment. Sensory information (internal and external) is processed in the Brain by using mindful and unconscious processing. Conscious processing produces the subjective sensation of 'consciousness' and includes the simultaneous activation of awesome neurophysiological systems. These structures are the Alerting System, the two Awareness Systems, the Affect and Arousal System, and the Attentional Systems. The purpose of awareness is to perceive threats or urgencies in the inner (inside the body) and exterior (outside the body) environments, and to mobilize the body's sources to deal with them.

Attention is usually classified into two awesome functions: bottom-up (or exogenous) attention, an externally brought on manner in which facts to be processed is chosen mechanically due to the fact of noticeably substantial points of stimuli; and top- down (or endogenous) attention, an internally triggered method in which facts is actively sought out in the environment primarily based on voluntarily chosen factors.

Bottom-up visual attention starts off evolved with simple visual processing alongside the visual cortical pathways. From the primary visual cortex (V1), feed-forward signals ascend to a couple of cortical areas and department into two important visible pathways: a ventral pathway dealing with object- and feature-based visible processes, and a dorsal pathway dealing with spatial- and movement-related visible processes.

Top-down visible attention is a voluntary method in which a unique location, feature, or object relevant to modern-day behavioural dreams is chosen internally and focused upon or examined. The essential effect of top-down interest is that neural endeavour is more desirable for the precise location/feature/object of interest compared to behaviourally beside the point stimuli, while suppression of neuronal responses is observed for beside the point stimuli. These kinds of response modulation have been located in areas of each ventral and dorsal visual pathways, such as V1, V2, V4, IT , MT, the PPC, and the PFC.

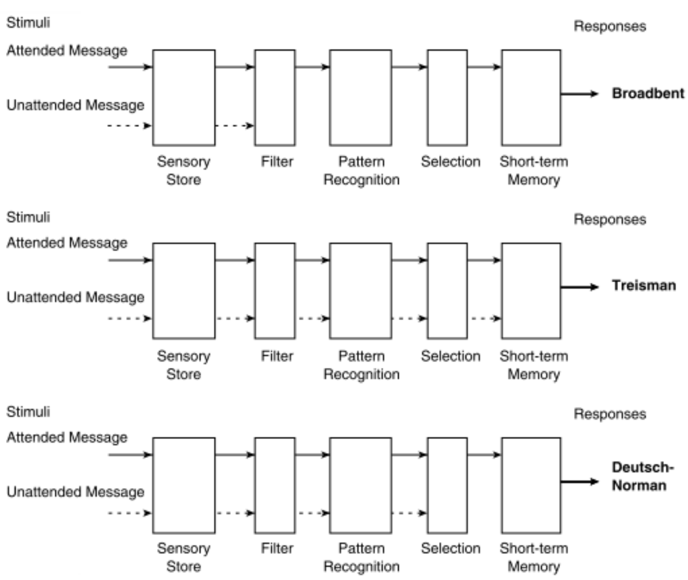
Multitasking may be defined as follows:

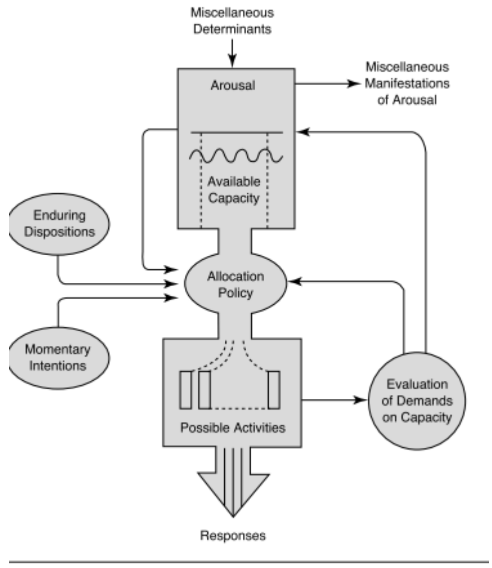
- It can mean performing two or more tasks simultaneously
- It can also involve switching back and forth from one thing to another
- Multitasking can also involve performing a number of tasks in rapid succession.

In the brain, multitasking is managed by mental executive functions. These executive functions control and manage other cognitive processes and determine how, when and in what order certain tasks are performed.

According to researchers Meyer, Evans, and Rubinstein, there are two stages to the executive control process.

1. The first stage is known as "goal shifting" (deciding to do one thing instead of another).
2. The second is known as "role activation" (changing from the rules for the previous task to rules for the new task).





Kahneman's capacity model of attention

Switching between these may only add a time cost of just a few tenths of a second, but this can start to add up when people begin switching back and forth repeatedly. Spreading attention so thin and constantly being distracted by different streams of information might have a serious, long-term, negative impact on how these connections form. Although the present findings do not demonstrate any causal effect, they highlight an interesting possibility of the effect of media multitasking on certain cognitive abilities, multisensory integration in particular. Media multitasking may not always be a bad thing

Q 2) What is the EEG? Brief the use of EEG in brain state (e.g. stress, work load, and attention) identification.

Electroencephalography , or EEG, is the physiological method of choice to record the electrical activity generated by the brain via electrodes placed on the scalp surface. For faster application, electrodes are mounted in elastic caps similar to bathing caps, ensuring that the data can be collected from identical scalp positions across all respondents. The function of EEG is:

- measure electrical activity generated by the synchronized activity of thousands of neurons (in volts).
- provide excellent time resolution, allowing you to detect activity within cortical areas -even at sub-second timescales.

As the voltage fluctuations measured at the electrodes are very small, the recorded data is digitized and sent to an amplifier. The amplified data can then be displayed as a sequence of voltage values. EEG is one of the fastest imaging techniques available as it often has a high sampling rate. One hundred years ago the time course of an EEG was plotted on paper. Current systems digitally display the data as a continuous flow of voltages on a screen.

As EEG monitors the time course of electrical activity generated by the brain, you can interpret which areas of the cortex are responsible for processing information at a given time:

1. Occipital cortex

   This part of the brain is primarily responsible for processing visual information. EEG experiments with visual stimuli (videos, images) often focus on effects in occipital regions.

2. Parietal cortex

   Parietal cortex is primarily responsible for motor functions and is active during self-referential tasks – when we are encountering objects or information that is important to us, for example.

3. Temporal cortex

   Temporal cortex has lateral aspects which are responsible for language processing and speech production. Medial (inner) regions are more active during spatial navigation.

4. Frontal cortex

   The frontal part of the human brain is enlarged compared to most other mammals. Basically, the frontal cortex is all about executive function: it helps us maintain control, plan for the future, and monitor our behaviour. Apart from the regional characteristics of where certain electrical activity originates, you can also analyse which frequencies primarily drive the ongoing activity.

Whenever your brain is in a certain state, the frequency patterns change, giving insight into cognitive processes.

- Delta (1 – 4 Hz) – Delta waves are examined to assess the depth of sleep. The stronger the delta rhythm, the deeper the sleep. Increased delta power (an increased quantity of delta wave recordings) has also been found to be associated with increased concentration on internal working memory tasks.
- Theta (4 – 7 Hz) – theta is associated with a wide range of cognitive processing such as memory encoding and retrieval as well as cognitive workload. Whenever we're confronted with difficult tasks (counting backwards from 100 in steps of 7, or when recalling the way home from work, for example), theta waves become more prominent. Theta is also associated with increased fatigue levels.
- Alpha (7 – 12 Hz) – whenever we close our eyes and bring ourselves into a calm state, alpha waves take over. Alpha levels are increased when in a state of relaxed wakefulness. Biofeedback training often uses alpha waves to monitor relaxation. They are also linked to inhibition and attention.
- Beta (12 – 30 Hz)- over motor regions, beta frequencies become stronger as we plan or execute movements of any body part. Interestingly, this increase in beta is also noticeable as we observe bodily movements of other people. Our brain seemingly mimics their limb movements, indicating that there is an intricate "mirror neuron system" in our brain which is potentially coordinated by beta frequencies.

- Gamma (>30 Hz, typically 40 Hz) – Some researchers argue that gamma reflects attentive focusing and serves as carrier frequency to facilitate data exchange between brain regions. Others associate gamma with rapid eye movements, so-called micro-saccades, which are considered integral parts for sensory processing and information uptake.

Analyzing EEG data can get quite challenging. Signal processing, artefact detection and attenuation, feature extraction, and computation of mental metrics such as workload, engagement, drowsiness, or alertness all require a certain level of expertise and experience to properly identify and extract valuable information from the collected data.

Q 3) Define various eye movement matrices and its possible relation with various cognitive processes. Brief the use of eye tracking in HCI.

Eye movements are a behaviour that can be measured and their measurement provides a sensitive means of learning about cognitive and visual processing. Although eye movements have been examined for some time, it has only been in the last few decades that their measurement has led to important discoveries about psychological processes that occur during such tasks as reading, visual search, and scene perception.

At one time, researchers believed that the eyes and the mind were not tightly linked during information processing tasks like reading, visual search, and scene perception. This conclusion was based on the relatively long latencies of eye movements (or reaction time of the eyes) and the large variability in the fixation time measures.

They questioned the influence of cognitive factors on fixations given that eye movement latency was so long and the fixation times were so variable. It seemed unlikely that cognitive factors could influence fixation times from fixation to fixation.

Actually, an underlying assumption was that everything proceeded in a serial fashion and that cognitive processes could not influence anything except very late in a fixation, if at all. However, a great deal of research using new eye trackers that enable better eye tracking has since established a tight link between the eye and the mind, and it is now clear that saccades can be programmed in parallel and, furthermore, that information processing continues in parallel with saccade programming.

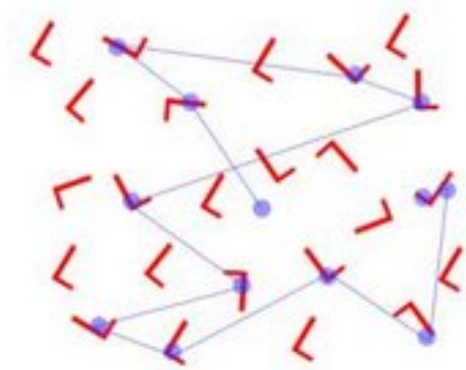
Task	Typical mean fixation duration (ms)	Mean Saccade Size (degrees)
Silent Reading	225-250	2 (8-9 letter spaces)
Oral Reading	275-325	1.5 (6-7 letter spaces)
Scene Perception	260-330	4
Visual Search	180-275	3

Fixation durations in search tend to be highly variable. Some studies report average fixation times as short as 180 ms while others report averages on the order of 275 ms. This is undoubtedly due to the fact that the difficulty level of a search (i.e., how dense or cluttered the array is) and the exact nature of the search task will strongly influence how long viewers pause on each item (see Figure 4 and Figure 5).

Typically, saccade size is a bit larger than in reading (though saccades can be quite short with very dense arrays).

When an array is very cluttered (with many objects and distractors), the search becomes more demanding than when the array is simple. The eye movements on each of these types of arrays typically reflect this property of an array. As the array becomes more complicated, we see an increase in the fixation duration and the number of fixations, as well as a decrease in the average saccade length.

The study of eye movements pre-dates the widespread use of computers by almost 100 years. Beyond mere visual observation, initial methods for tracking the location of eye fixations were quite invasive — involving direct mechanical contact with the cornea. Dodge and Cline (1901) developed the first precise, non-invasive eye tracking technique, using light reflected from the cornea. Their system recorded only horizontal eye position onto a falling photographic plate and required the participant's head to be motionless. In more recent times, eye tracking in human-computer interaction has shown modest growth both as a means of studying the usability of computer interfaces and as a means of interacting with the computer. As technological advances such as the Internet, e-mail, and videoconferencing evolved into viable means of information sharing during the 1990s and beyond, researchers again turned to eye tracking to answer questions about usability and to serve as a computer input device.



Q 4) Describe the visual information processing the brain? Brief on visual perception.

The axons of ganglion cells exit the retina to form the optic nerve, which travels to two places: the thalamus (specifically, the lateral geniculate nucleus, or LGN) and the superior colliculus. The LGN is the main relay for visual information from the retina to reach the cortex. Despite this, the retina only makes up about 20% of all inputs to the LGN, with the rest coming from the brainstem and the cortex. So more than simply acting as a basic relay for visual input from retina to cortex, the LGN is actually the first part of our visual pathway that can be modified by mental states.

The superior colliculus helps us to control where our head and eyes move, and so determines where we direct our gaze. Saccades, the jumpy eye movements that you are using as you read this text, are also controlled by the superior colliculus. As with the LGN, the superior colliculus receives strong input from the cortex, which provides the dominant command as to where our gaze moves.

From the thalamus, visual input travels to the visual cortex, located at the rear of our brains. The visual cortex is one of the most-studied parts of the mammalian brain, and it is here that the elementary building blocks of our vision – detection of contrast, colour and movement – are combined to produce our rich and complete visual perception.

Most researchers believe that visual processing in the cortex occurs through two distinct 'streams' of information. One stream, sometimes called the What Pathway, is involved in recognising and identifying objects. The other stream, sometimes called the Where Pathway, concerns object movement and location, and so is important for visually guided behaviour.

Our visual cortex is not uniform, and can be divided into a number of distinct subregions. These subregions are arranged hierarchically, with simple visual features represented in 'lower' areas and more complex features represented in 'higher' areas.

At the bottom of the hierarchy is the primary visual cortex, or V1. This is the part of visual cortex that receives input the thalamus. Neurons in V1 are sensitive to very basic visual signals, like the orientation of a bar or the direction in which a stimulus is moving. In humans and cats (but not rodents), neurons sensitive to the same orientation are located in columns that span the entire thickness of the cortex.

That is, all neurons within a column would respond to a horizontal (but not vertical or oblique) bar. In a neighbouring column, all neurons would respond to oblique but not horizontal or vertical bars (see image below). As well as this selectivity for orientation, neurons throughout most of V1 respond only to input from one of our two eyes. These neurons are also arranged in columns, although they are distinct from the orientation columns.

Moving up the visual hierarchy, neurons represent more complex visual features. For example, in V2, the next area up in the hierarchy, neurons respond to contours, textures, and the location of something in either the foreground or background. Beyond V1 and V2, the pathways carrying What and Where information split into distinct brain regions. At the top of the What hierarchy is inferior temporal (IT) cortex, which represents complete objects – there is even a part of IT, called the fusiform face area, which specifically responds to faces. The top regions in the Where stream are involved in tasks like guiding eye movements (saccades) using working memory, and integrating our vision with our body position (e.g. as you reach for an object).

In summary, the visual cortex shows a clear hierarchical arrangement. In lower areas (those closest to incoming light, like V1), neurons respond to simple visual features. As the visual input works its way up the hierarchy, these simple features are combined to create more complex features, until at the top of the hierarchy, neurons can represent complete visual objects such as a face.

Q 5) Define various emotional states (valence and arousal model). How to identify and measure the emotion from a subject? (facial expression, EEG, ECG, GSR, speech etc.)

Researchers wanted to see how different emotions were related. They needed a model-type that would show unique relationships within a visual framework. They had already constructed linear models which displayed how different sets of emotions were related along a continuum (the same concept as a number line). However, these models weren't able to adequately show how the entire range of emotions are related. Instead, they used a circular model that was divided into quadrants with axes of crossed continua.

This type of model serves three purposes:

- It is a pictorial representation of a range of data.
- It provides a field on which the relationship between different variables can be seen.
- The items can be equally spaced to show uniformity and precision.

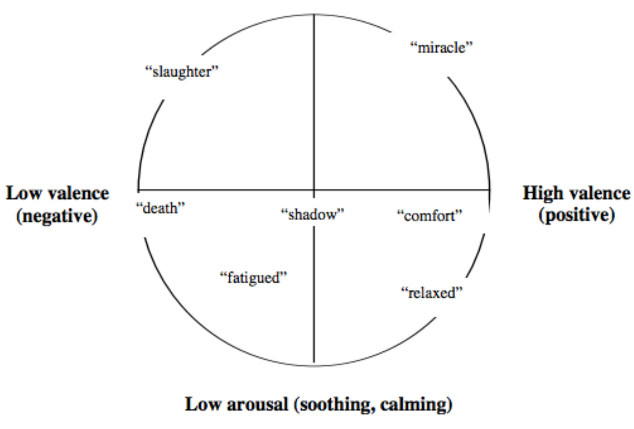
In general, emotional experiences can be described by two terms:

- Valence
- Arousal

Though the two terms are often confused, the difference is pretty simple. Valence is positive or negative affectivity, whereas arousal measures how calming or exciting the information is. Arousal comes from our reptilian brain. It inspires a fight-or-flight response, that, evolutionary, aided our survival. They're quite different in definition, and some studies have shown that our brain processes each factor independently. One study manipulated odour valence (pleasant or unpleasant) as well as intensity, and they found that odour intensity triggered amygdala activity, where valence triggered orbitofrontal activity.

The common framework for dealing with emotional experience is characterized in a two-dimensional space. Valence ranges from highly negative to highly positive, and arousal ranges from calming/soothing to exciting/agitating.

Figure 1. Affective experiences may be best described in two dimensions: Valence refers to how positive or negative an event is, and arousal reflects whether an event is exciting/agitating or calming/soothing. Words have been placed at locations within this space, indicating their approximate valence and arousal ratings (ratings from Bradley & Lang, 1999 [#11]).



High-arousal emotional events cause you to hone in on the arousing stimuli. They're better encoded and better recalled than non-arousing events. Emotional arousal works as a sort of blinder to other neutral stimuli. A study showed that, instead of increasing overall attention to an event, an emotionally arousing stimulus decreased attentional resources available for information processing and focused attention only on the arousal-eliciting stimulus."

Studies show that either positive or negative-valence emotions, the emotions at the extremities, are remembered better. For example, emotional words are remembered better than neutral words.

However, much of the mechanism behind the stronger memory of positive or negative valence emotion is elaborating (assigning meaningful information to what you're trying to remember), which takes two forms:

- Autobiographical
- Semantic

There are various ways for measuring emotional states:

- Self-report measure of emotion
- Autonomic measure of emotion
- Startle response magnitude as a measure of emotion
- Brain state as a measure of emotion.
    - Electroencephalography
    - Neuroimaging studies
- Behaviour as a measure of emotion
    - Vocal characteristics
    - Facial observation
        - Observer ratings
        - Electromyography
    - Whole body behaviour

Different measures of emotion appear sensitive to different dimensional aspects of state (e.g., facial EMG is sensitive to valence, whereas skin conductance is sensitive to arousal) and are not strongly related to one another. Practically speaking, then, there is no "gold standard" measure of emotional responding. For theories of emotion, this means that there is no "thing" that defines emotion, but rather that emotions are constituted by multiple, situationally and individually variable processes.