Chapter 3
How to Measure Attention?

Matei Mancas and Vincent P. Ferrera

Researchers who are interested in attention generally have one or more of the following goals: (1) to identify sources of information in the environment that are selected and prioritized by the observer, (2) to quantify the effect of attention on task performance, and (3) to identify physiological or neural correlates of attention. When considering methods to measure attention, it is important to distinguish between overt and covert orienting mechanisms. Overt attention is expressed by movements of the body and can be measured directly by determining the position and velocity of the relevant effectors—primarily the eyes but also head or pointers (hand, mouse, etc.). Covert orienting refers to the ability to direct attention without body movement and is primarily measured by differences in task performance (e.g., reaction time, detectability, or discriminability) that cannot be attributed to changes in the external stimulus.

In this chapter we will focus on techniques that provide fine-grain spatial and temporal information about attentive responses at a macro scale. We do not discuss the many psychophysical paradigms that have been used to infer attention based on the speed and accuracy of observer judgments. Measurements of a single or multiple neurons using micro-electrodes are not described here either. However, the chapters of the second part of the book (Chaps. 4, 5, and 6) deal also with micro measurements.

At a macro scale, the attentive response can be either measured directly in the brain or indirectly through participants' behavior. Only one of the techniques that is described here is based on participant active feedback: mouse tracking. This is

M. Mancas (✉)
Numediart Institute, University of Mons, Mons, Belgium
e-mail: matei.mancas@umons.ac.be

V. P. Ferrera
Zuckerman Institute on Mind Brain and Behavior, Columbia University, New York, NY, USA

M. Mancas et al. (eds.), *From Human Attention to Computational Attention*,
https://doi.org/10.1007/978-3-031-84300-6_3

because mouse tracking feedback is an approximation of eye tracking and it is an emerging approach of interest for the future: it requires less time and less money and provides more data than classical eye tracking. All the other methods are direct or indirect and provide objective measures of attention. This chapter first describes the indirect methods, while direct methods are mainly dealt with in the second part.

3.1 Indirect Measures of Attention

3.1.1 Eye Tracking: A Gold Standard for Overt Attention

If "the eyes are windows to the soul," eye tracking consists of taking a look at it. Indeed, eye tracking is probably the most widely used tool for measuring visual attention. Although attention can be directed without moving the eyes, it is generally the case that humans look where they attend and vice versa. There is ample neurophysiological support for this proposition as several structures that are involved in attention—in prefrontal cortex, parietal cortex, and the midbrain [1]—are also involved in guiding voluntary eye movements.

Eye trackers are devices that determine the orientation of the eye relative to the head (eye-in-head) or to an external frame of reference (eye-in-space). If head position is known, then the orbital position of the eye (eye-in-head) is sufficient to determine gaze direction (eye-in-space).

Eye-tracking technology has evolved over time. Different technologies are described in [2]. The first eye tracker is attributed to Edmund Huey in 1898, while the first eye movement recordings were made in 1937 by Buswell [3]. Eye trackers have been built using optical, mechanical, and electromagnetic principles.

One of the earliest techniques to be widely used is EOG (electro-oculography). The eye itself generates an electric dipole oriented along the corneo-retinal axis. This potential can be measured by placing electrodes on the skin around the eye. From these electrodes, the eye orientation relative to the head can be reconstructed. To determine the orientation of the eye in space, the head must either be attached to a fixed system (chin rest or bite bar) or a head tracking system must be used in addition to the EOG. EOG signals are noisy and confounded by skin conductance and the activity of facial muscles. Reliable measurements typically require averaging over many trials.

A more precise method was developed in the 1960s [4, 5] using the scleral search coil. Here, a loop of wire is embedded in an annular contact lens placed around the cornea. A small electric current is passed through the wire, generating a magnetic dipole whose orientation moves with the eye. The subject sits with their head inside an oscillating magnetic field generated by a pair of large (roughly 2–3 feet in diameter) field coils. Electronics are used to sense the orientation of the scleral coil and hence the orientation of the eye. This system measures eye orientation relative to the field coils, which are fixed in space. The head generally needs to be stabilized to avoid confusing the rotation of the eye with translations due to head movement. A separate head coil can be used to record head movement. Binocular search coil systems allow experimenters to reconstruct vergence angles. Torsional eye movements can also be recorded. Scleral search coil systems provide continuous temporal output, limited only by front-end filtering and the sampling rate of the recording device used to convert the analog signal to digital samples. Spatial resolution is typically 0.1 deg. of visual angle or better and noise is extremely low. Contact lens search coils can only be worn for a short time (<30 min) as they cause an increase in intraocular pressure during the time that they are in contact with the sclera. This method should be used only under the supervision of a trained clinical ophthalmologist.

The technique that most commercial and research solutions use is video oculography (VOG), based on a video camera to detect the pupil and corneal reflection. An infrared light source illuminates the eyes. The light is either reflected (bright pupil) or absorbed (dark pupil) by the pupil and image processing software (usually embedded in dedicated hardware) is used to detect the edges of the pupil either by filling in or fitting an ellipse to the edge of the iris (Fig. 3.1). This processing also provides an estimate of pupil size. Crosshairs identify the horizontal and vertical position of the center of the pupil. Some light is also reflected from the cornea and is called the corneal reflection (CR). The position of the pupil and corneal reflection is sensitive to head movement. However, the difference (pupil—CR) discounts the influence of head motion and gives a robust estimate of eye orientation in space. Nevertheless, for precise measurements, it is best to stabilize the head with a chin rest or bite bar.

It must be kept in mind that VOG trackers operate on a two-dimensional image of the eye. To obtain eye orientation, the appropriate transformation must be done considering the geometry of the camera relative to the eye and the projection of a 3D sphere onto a 2D image. Alternatively, a look-up table matching eye position to tracker output can be generated by having subjects fixate on targets at known positions. A grid of at least nine positions should be used for this calibration. VOG

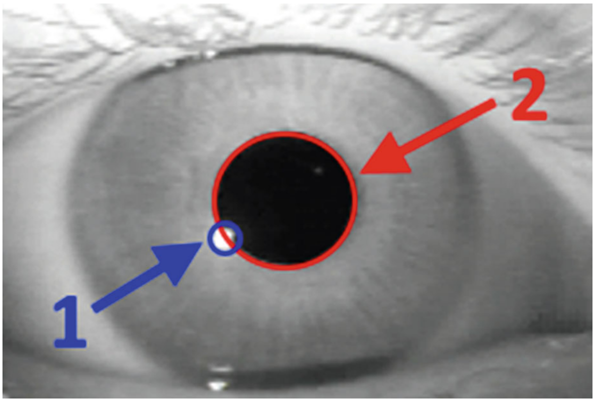

Fig. 3.1 The relative position of the pupil (arrow 2) and the corneal reflection (arrow1) are used to compute the gaze direction

systems work best when the optics of the camera are aligned with the optical axis of the eye when the subject is looking straight ahead (primary position). An infrared or "hot" mirror placed in front of the eye can be used to achieve this alignment. The infrared mirror is transparent to visible light. This way, the subject can directly view the visual display or scene through the mirror, while the camera is placed off to the side.

The temporal resolution of VOG systems is limited by the framerate of the camera and the speed of the image processing algorithm that identifies the pupil and corneal reflection. Commercially available systems range from 30 Hz to over 1000 Hz. Spatial resolution is limited by the resolution of the camera. Typically, this is enhanced by using telephoto and close-up lenses to magnify the image of the eye. Many systems provide spatial resolution comparable to search coils (0.5 deg. of visual angle or less). Drawbacks of VOG systems include sensitivity to stray light, which may cause large apparent changes in eye position. Furthermore, these systems are unable to function when the subject blinks, and typically set their output to a default value whenever this happens.

While the fundamental technique is most of the time the same, the physical apparatus of the eye tracker can be very different. The main eye-tracking manufacturers such Tobii [6] or SR Research [7] propose the system under different forms. SMI eye tracking is another historical eye-tracker manufacturer which was discontinued but Gaze Intelligence [8] provides support and access to the product manuals and software for all SMI solutions.

Recently, VOG approaches have started to use deep learning techniques. Indeed, traditional algorithms implement the image to direction transformation using logical rules and calibration data as we discussed in the previous sections, while deep learning led to data-driven approaches where the algorithm learns the nonlinear relationship between eye direction and images of the eye. In this case the camera does not need to be fully in front of the eye (it might be a little on the side) which is very practical especially for glasses-like hardware where cameras in front of the eye decrease the field of view of the user. The AI-based deep learning approach also does not need a specific well done calibration, it might even work without calibration which is especially interesting in ecological hardware like eye-tracking glasses used during sport, for example, where head movements introduce changes in the glasses position with respect to the eyes. In this case a classical approach would need to re-calibrate which is not possible, while an AI-based approach will be able to recover with no problem from the glasses/eye sudden changes in relative position. Issues in AI-based models are of course the need of huge image datasets with the same view of the eye and very different illumination, eye, and skin color. Pupil Labs [9] is pioneering AI-based eye tracking for commercial use.

Finally, there are two eye-tracking approaches which are just at their inception but have probably an interesting future. The first one uses dynamic vision sensors (DVS) which are cameras that only send events (i.e., pixels which change between the frames). Those cameras substantially reduce the amount of video data and their application to pupil changes and eye tracking leads to an impressive sampling rate of 10,000 Hz with good accuracy for a wearable device (between 0.45° and 1.75°

of angle) [10]. While DVS is still experimental, another approach which is not based on cameras is more developed. This approach is based on the micro-electro-mechanical system (MEMS). The MEMS scanner emits infrared rays toward the cornea at a glancing angle with a high frequency (500 Hz for the commercial AdHawk MindLink [11]). The infrared rays pass through a system ensuring even coverage of the corneal area. The photodiodes located in the eyeglass frame detect the reflected infrared light from the cornea. Obviously eye movements imply cornea motion which will modify the reflected infrared light and this modification lets the system compute the gaze direction. This eye-tracking system does not require cameras and image processing. The consumption needs are very small compared to classical cameras, the frequency is higher, and the accuracy is already close to existing commercial glasses [12] from Tobii, SMI, or Pupil Labs.

Eye tracking devices can be very diverse depending on the applications. First, we can cite the very precise systems, mainly used in psychology setups such as the SR Research Eyelink [7] which aim in getting the best data with small errors (down to 0.15° error) and very high sampling rate (up to 2000 Hz). Devices can synchronize with a large variety of other research devices such as EEG (Fig. 3.2, left) or even MEG (Fig. 3.2, middle) and MRI (Fig. 3.2, right). They can be used with a chin rest device to avoid head movements for even better accuracy (less than 0.15°). Some versions can also work with animals for experiments on primates, for example.

Some eye trackers are directly incorporated into the screen (Fig. 3.3, left) which is used to present the data. This setup has the advantage of a very short calibration, but it can only be used with its own screen. Those systems have largely been replaced by systems which have a screen adapted to the eye tracker but which can be removed and put on another screen as the one in Fig. 3.3 right. Resolutions here remain high with high accuracy (0.5° of angle or less) and high sampling rates.

Wearable eye-tracking glasses (Fig. 3.4) can be used in very ecological setups, even outside on real-life scenes. An issue of those systems is that it is not easy to aggregate the data from several viewers as the scene which is viewed is not the same. The aggregation needs a non-trivial registration of the scenes which might require installing markers before the experiment or semi-automatic registration partially made by a registration algorithm but which needs human validation. New models are more and more discrete such as Tobii glasses 2 (Fig. 3.4, middle) or Pupil Labs glasses (Fig. 3.4, right) compared to older SMI systems (Fig. 3.4, left).

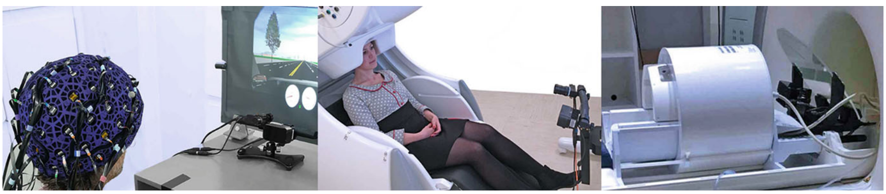

Fig. 3.2 SR Research Eyelink [7] example of precise research-purpose system. Eye tracker used along with (1) EEG (left), (2) MEG (middle), and MRI (right). (Images adapted from [7] website)

Fig. 3.3 Example of eye-tracking device included in a high-resolution screen (left) or attached to it (right) [6]. (Image adapted from Tobii website [6])

Fig. 3.4 Eye tracking embedded in glasses. Most of the time a simple USB cable goes into a small device or smartphone which can be stored into the user pocket. Left, you can see an older SMI system [8]; middle, a Tobii system [6]; and right, a Pupil Labs system [9]. (Image adapted from [6, 8, 9])

The incorporation of a forward-looking scene camera allows eye movements to be registered with the changing visual scene in real time.

Eye trackers are also increasingly present in augmented reality (AR) and virtual reality (VR) devices. While Microsoft was a leader in the high-quality AR market with Hololens 2 [13] (Fig. 3.5, left) which includes eye tracking, the withdrawal of Hololens 3 among other signs shows that the company might invest less in the direction of AR. In the field of high-quality AR Apple presented its Apple Vision Pro [14] (Fig. 3.5, right) which also provides an integrated eye tracker. This eye tracker is used for user interface interaction along with hand interaction and foveated vision allowing high frequency and high resolution where the user eyes direct. While information about the raw data of the eye tracking on this device are still scarce, studies show an accuracy around 1° of angle [15].

Several VR headsets also include eye tracking. The Meta Quest Pro [16] (Fig. 3.6, left), for example, includes an eye tracker with an accuracy average around 1.6° of angle [17]. However, it is not certain that there will be a Meta Quest Pro 2 and the classical (not pro) version has no eye tracking inside. The PICO family VR headsets such as the Neo3 Pro Eye, PICO 4 Pro, and PICO 4 Enterprise (Fig. 3.6, right) also provide built-in eye trackers [18].

Fig. 3.5 Eye tracking embedded in high-end AR devices. Left: Hololens 2 [13]. Right: Apple Vision Pro [14]. (Image adapted from websites of [13, 14])

Fig. 3.6 Eye tracking embedded in VR devices. Left: Meta Quest Pro [16]. Right: PICO 4 Enterprise [19]. (Image adapted from websites of [16, 19])

While it is also possible to find a built-in eye tracking on the HTC Vive Pro Eye [20], HTC now uses an add-on for full-face tracking (including eye tracking) [20] (Fig. 3.7, left). Other manufacturers such as Pupil Labs provide “add-ons” which can fit in several VR headsets [9] (Fig. 3.7, middle). Some add-ons are even using the novel MEMS technology such as AdHawk [11] or Inseye [21] (Fig. 3.7, right). Those devices are used in several applications such as VR games and interactions or professional training.

Inexpensive devices (Fig. 3.8) appeared some years ago with a price as low as 100 EUR [22] which is a fraction of the price of a professional eye tracker. However, the manufacturer, Eye Tribe, was bought by Oculus. Tobii eye tracker 5 [23] can be found around 300 EUR, the HTC Vive face tracker around 250 EUR [20], or the Kexxu Eye tracking glasses [24] at less than 800 EUR. An issue with these eye trackers is that they are often packaged with minimal software and it is often difficult to synchronize the stimuli and the related eye movement data. These eye trackers are mostly used as real-time human-machine interaction devices in gaming

Fig. 3.7 Eye-tracking "add-ons" which can be added to existing VR devices. Left: Full-face tracking of HTC [20]. Middle: Pupil Labs for PICO Neo [9]. Right: MEMS-based eye tracking from Inseye [21]. (Image adapted from websites of [9, 20], and [21])

Fig. 3.8 Low-cost eye-tracking devices. Left: the former Eye Tribe device [22]. Middle: Tobii eye tracker 5 [23]. Right: Kexxu Eye tracking glasses [24]. (Image adapted from websites of [22, 23], and [24])

applications. However, it is possible to use their data and there are even open source projects that allow recording of data from low cost eye trackers like Ogama [25], but mainly on still images and not moving stimuli.

Finally, webcam-based software is freely available such as OpenTrack [26], GazeRecorder [27], Beam [28], or even Javascript for any browser with WebGazer.js [29]. The webcam from a computer can be used, but also sometimes a smartphone camera.

Eye movement behavior has a rich variety of features that are indicative of attention. In primates, voluntary eye movements consist of saccades (rapid changes in position with peak angular velocity >100 deg./sec first described by French ophthalmologist Javal in 1879 [3]), vergence (changes in the alignment of the two eyes), and smooth pursuit (slow movements, generally under 100 deg./sec, that track small moving targets). Between these movements are periods of fixation, though microscopic movements (drift, tremor, and microsaccades) may still occur even when the eye is relatively still. Fixations can be detected using clustering algorithms [30, 39] or simply by using a double threshold: a time threshold and a spatial threshold to be sure that the gaze remains focused in a small region.

Fixation duration can be a measure of attention and awareness [31, 32]. Fixations and intervening movements can be used to generate scanpaths (Fig. 3.9) or heatmaps (Fig. 3.10). A heatmap is a low-pass filtered accumulation of scan paths and it indicates the average attention attraction of each pixel. Usually for a result to be significant there is a need of a minimum of 10 participants per stimulus.

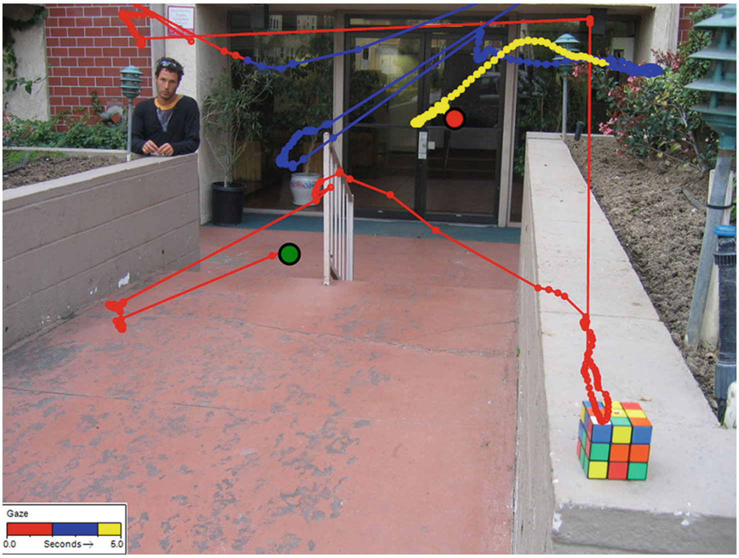

Fig. 3.9 Example of eye scanpath provided by eye-tracking systems

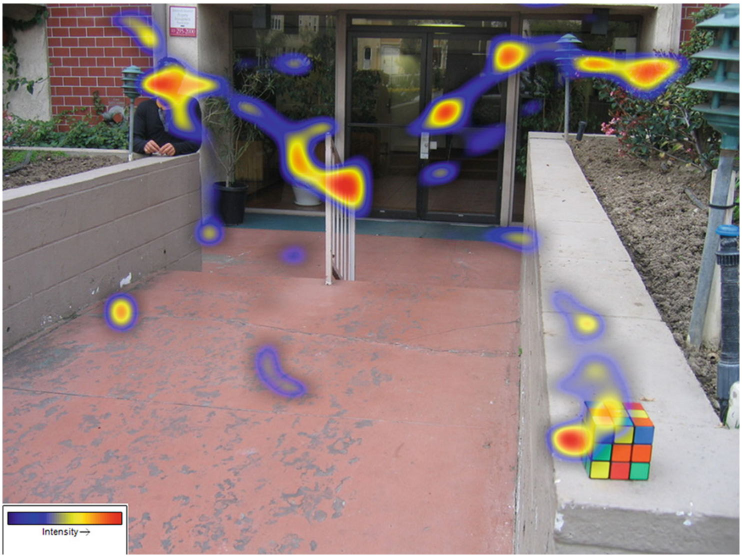

Fig. 3.10 Example of attention heatmap averaged over participants

During fixations, subjects often make very small eye movements called microsaccades [33]. These are saccades with amplitudes of less than 2 degrees of visual angle. Spontaneous microsaccades are often correlated with attention [34].

When viewing static scenes at a fixed depth, the most common eye movements are saccades, which normally occur roughly 2–3 times per second. The onset of a saccade can be detected to within a few milliseconds using algorithms based on eye velocity or acceleration. The latency of saccades relative to the sudden appearance of a target is generally 150–300 msec. Variations in saccade latency may be related to attention [35]. Attention may alter saccade direction [36] or may result in curved saccade trajectories [37]. Recent studies [32] show that awareness about an object in the scene is (1) sudden and not linear which fits with the idea that awareness is linked to a sudden synchronization of neurons [38] and (2) it might take place in two steps, the first one as early as 100 ms detected around 200 ms by early negativity for object detection and the second is more about identification and relation to the task which begins around 200 ms and can be observed around 300 ms with P300 signal.

3.1.2 Mouse Tracking: The Low-Cost Eye Tracking

Although eye tracking may be the most reliable ground truth in the study of overt visual attention, it has several drawbacks in addition to the high cost of the professional devices:

- It needs minimal practice for the operator.
- The user's head might need to be stabilized.
- The calibration process might be long.
- The infrared light pointing the eyes might induce eye fatigue and dryness especially during long tests.
- The system might work much less well depending on the user eye color or if they wear glasses.

A much simpler way to acquire data about visual attention may be the use of mouse tracking. The mouse can be precisely followed while an Internet browser is open by using a client-side language like JavaScript. The mouse's precise position on the screen can be either captured using homemade code or existing libraries like [40, 41] or commercial solutions such as [42]. This technique may appear as not very reliable; however, its accuracy depends on the context of the experiment.

The first case is the one where the participant is unaware of the fact that the mouse motion is recorded. In this case mouse motion is not accurate enough. Indeed there is no automatic following of the eye gaze by the hand even if a tendency of the hand (and consequently the mouse) to follow the gaze is visible. Sometimes the mouse is only used to scroll a page and the eyes are very far from the mouse pointer, for example.

The second case is the one where the participant is aware of the experiment and has a task to follow. This can go from a simple "point the mouse where you look" instruction as in [43] where mouse tracking was used for the first time for saliency evaluation to more recent approaches as the one of SALICON in [44] where multi-resolution interactive cursor mimicking the fovea resolution is used to encourage

Fig. 3.11 Left: initial presented image. Right: mouse tracking heatmap after averaging across participants

people to point the mouse cursor where they look. Indeed, as the image resolution is decreased far from the cursor, people tend to point at the locations they are interested in to have a full-resolution view of those regions.

In this second case where the participant is aware about his mouse motion tracking, the results of mouse tracking are very close to eye tracking [22, 44]. However, small or unconscious eye movements may be missed (Fig. 3.11).

The main advantages of mouse tracking are low price and the complete transparency for the users (they just move a mouse pointer). The output can be the same as in eye tracking. It can either be a heatmap (Fig. 3.9), but also scan paths, raw data, etc.

However, mouse tracking has also several drawbacks:

- The first place where the mouse pointer is located is quite important as the observer may look for the pointer. Should it be located outside the image or in the center of the image? Ideally, the pointer should initially appear randomly in the image to avoid introducing a bias of the initial position of the pointer.
- Mouse tracking only highlights areas that are consciously important for the observer. This is more a theoretical drawback than a practical one as one should try to predict the overtly interesting regions.
- The pointer hides the image region it overlaps; thus, the pointer position is never on the important areas but very close to them. This drawback may be partially eliminated by the low-pass filter step performed after the mean of the whole observer set. It is also possible to make a transparent pointer as in [44].

Mouse tracking was neglected with few publications since [43] and somehow considered as a “poor man’s eye tracking.” However, the rise of learning-based computational models using deep neural networks, which need huge datasets to provide correct results, has changed the situation. Mouse tracking can be done online by a virtually unlimited number of participants allowing the generation of big datasets of mouse tracking data. As eye tracking can only provide datasets with a limited number of stimuli and users per stimulus, even if they are more precise, the development of mouse tracking has certain advantages that complement eye tracking. Moreover, the combined use of eye and hand tracking can also provide insight into the deployment of attention in natural tasks [45].

3.2 Direct Measures of Attention

3.2.1 EEG: Get the Electric Activity from the Brain

The EEG technique (electroencephalography) uses electrodes placed on the participant's scalp. Those electrodes amplify electrical potentials originating in the brain. An issue of this technique is that the skull and scalp attenuate and distort those electrical signals introducing some noise.

While classical research setups have a high number of electrodes up to 256 (Fig. 3.12) with manufacturers like Neuroscan [46], Brain Products [47], or BioSemi [48], some low-cost commercial systems like NeuroSky [49], Muse [50], or Emotiv [51] are more compact with much less electrodes (16 to 1), easier to install and calibrate, and easier to wear in ecological situations (Fig. 3.13). While the latter are easier to use, they are obviously less precise.

EEG studies provided interesting results such as the modulation of the gamma band [52] during selective visual attention. Other papers [53] also provide cues about the alpha band modification during attentional shifts.

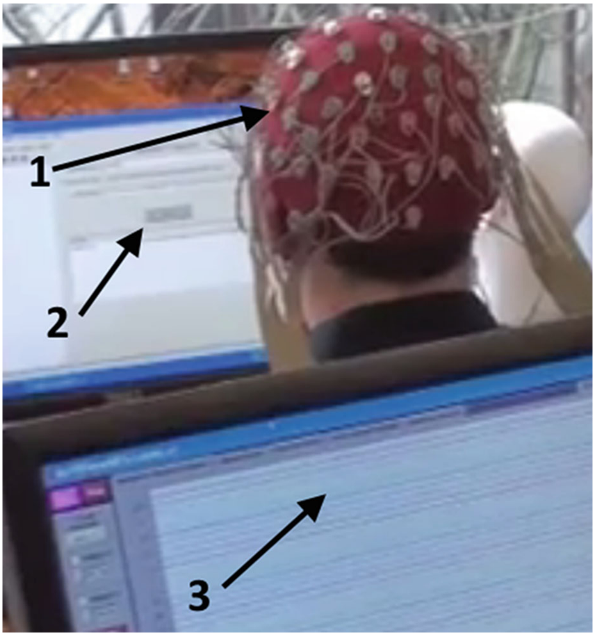

Fig. 3.12 Example of a research EEG device with a lot of electrodes (1), screen for the participant to visualize stimuli and tasks (2), and screen for the operator to visualize the signals (3)

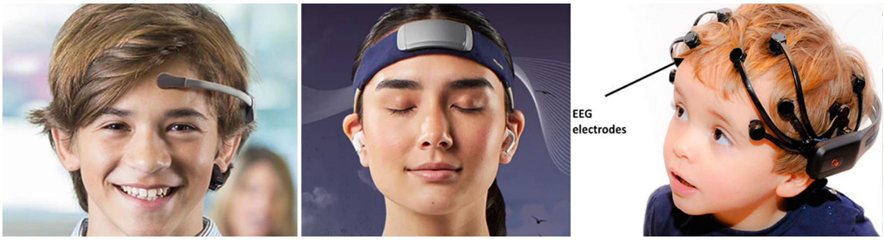

Fig. 3.13 A low-cost commercial EEG. Left: NeuroSky [49]. Middle: Muse [50]. Right: Emotiv [51]. Images adapted from websites of [49–51]

One very important cue about attention which can be measured using EEG is the P300 event-related potential (ERP).

The work of Näätänen et al. [54] in 1978 on auditory attention provided evidence that the evoked potential has an enhanced negative response when the subject was presented with rare stimuli compared to frequent ones. This negative component is called the mismatch negativity (MMN), and it was observed in several experiments. The MMN occurs 100–200 msec after the stimulus, a time that is perfectly in the range of the preattentive attention phase.

Depending on the experiments, different auditory features were isolated: audio frequency [55], audio intensity [56, 57], spatial origin [58], duration [59], and phonetic changes [60]. All these features were not salient alone, but saliency was induced by the rarity of each one of these features.

The study of the MMN signal for visual attention has been investigated several times in conjunction with audio attention [61–63]. But a few experiments were made using only visual stimuli. Crottaz-Herbette in her thesis [64] conducted an experiment in the same conditions as for auditory MMN and showed a strong increase of the negativity of the evoked potential when seeing rare stimuli compared with the evoked potential when viewing frequent stimuli. The visual MMN occurs from 120 to 200 msec seconds after the stimulus. The 200 msec frontier approximately matches the 200 msec needed to initiate a first eye movement, thus engaging the "attentive" serial attentional mechanism. As for the audio MMN detection, no specific task was presented to the subject who only had to hear the stimuli; this MMN component is thus preattentive, unconscious, and automatic. This study and others [65] also suggest the presence of a MMN response for the somatosensory modality (touch, taste, etc.) The MMN seems to be a universal component of the brain response reflecting an unconscious preattentive process. Any unknown stimulus (novel, rare) will be very salient as measured by P300. Rarity or novelty is a major driver of the attentional mechanism for visual, auditory, and all the other senses. Recent studies with only visual stimuli [32] show that the passage from unconscious preattentive process to first awareness needs an eye fixation lasting at least 100 ms (and is detected on the EEG after 200 ms) for a first object of interest awareness but without real identification of the object. From fixations lasting more than 200 ms (detected by P300), the object identification awareness can occur.

3.2.2 Functional Imaging: fMRI

MRI stands for magnetic resonance imaging. The main idea behind this kind of imaging system is that the human body is mainly made of water which is itself composed of hydrogen atoms that are composed of a single proton. Those protons have a magnetic moment (spin) which is randomly oriented most of the time. The MRI device uses a very high magnetic field (B0) to align the magnetic moment of a small fraction of protons in the patient's body, and a gradient field to control spin frequency. Radio frequency (RF) pulses are used to drive the proton spins into a plane orthogonal to B0. As the spins reorient or "relax" parallel to the orientation of B0, RF emissions are produced. Those emissions are captured and an inverse Fourier transform is used to construct an image where higher intensity levels mean that there are more protons, therefore more water in the body parts (e.g., like in fat), and darker gray levels reveal regions with less water (like bones).

MRI was initially an anatomical imaging technique, but it was soon discovered that the susceptibility artifact created by iron in the blood could be used to measure blood volume and oxygenation. Since blood volume and oxygenation respond to the metabolic demands of neural tissue, they can be used as a proxy for neuronal activity. In that way, when a region in the brain is activated, the blood may have an increased flow. The hemodynamic response has multiple components that bear a complicated relationship to the metabolic and electrical activity of the neural tissue. Nevertheless, fMRI imaging is capable of detecting the areas in the brain which are more or less active and has become a great tool for neuroscientists to visualize which area in the brain responds during an attention-related patient exercise (Fig. 3.14).

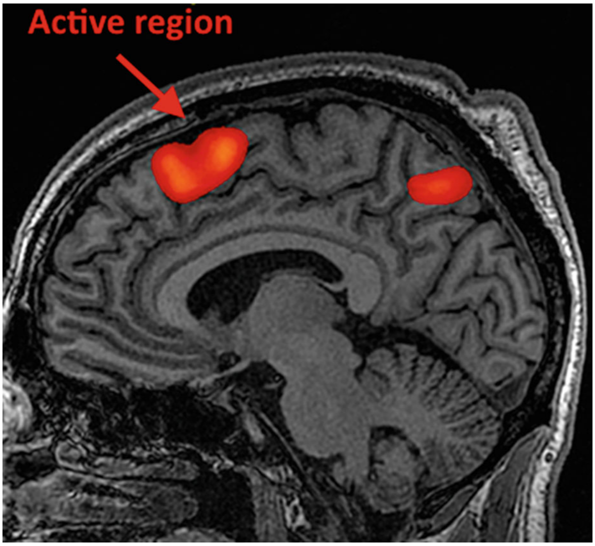

Fig. 3.14 Example of fMRI output: red active regions superimposed on an anatomical MRI sagittal image. (Adapted from [69])

3.2.3 Functional Imaging: MEG

MEG stands for magnetoencephalography. The idea is simple: while the EEG detects the electrical field which is heavily distorted when traversing the skull and skin, MEG detects the magnetic field induced by this electrical activity. The magnetic field has the advantage of not being influenced by the skin or the skull. While the idea is simple, in practice the magnetic field is very weak which makes it very difficult to measure. This is why the MEG imaging is relatively new: the technological advances that allow MEG to be effective are based on SQUID (superconducting quantum interference devices). The magnetic field of the brain can induce electricity in a superconducting device which can be precisely measured. Modern devices have spatial resolutions of 2–4 millimeters and temporal resolutions of about 10 milliseconds. Moreover, MEG images can be superimposed on MRI anatomic images which helps to rapidly localize the main active areas. Finally, participants in MEG imaging can have an upright seated position (Fig. 3.15) which is more natural during testing than the horizontal position of fMRI or PET scan.

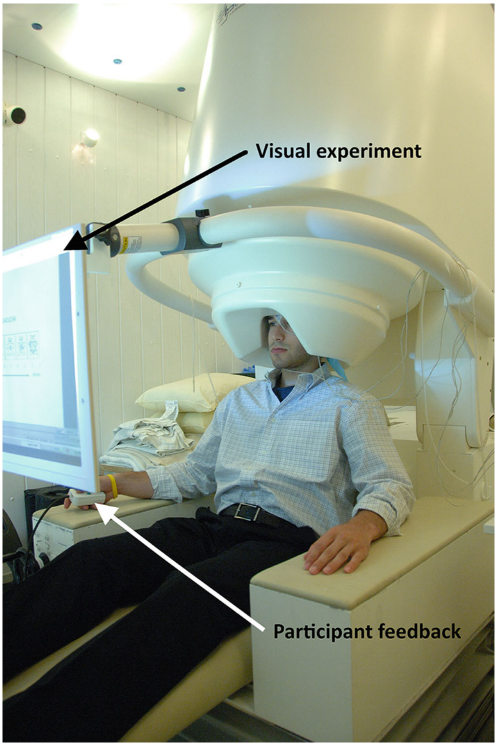

Fig. 3.15 A participant set into the MEG device and a visual experiment. Adapted from [66]

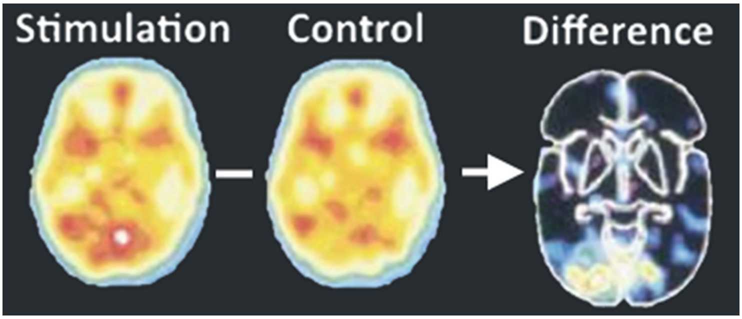

Fig. 3.16 Example of output in case of a repetitive visual pattern (flickering). The difference let us see the areas activated by the stimulus. (Adapted from [67])

3.2.4 Functional Imaging: PET Scan

Like fMRI and MEG, PET (positron electron tomography) scanning is also a functional imaging tool that produces a signal proportional to brain activity. The main idea of PET scan is that a mildly radioactive substance such as F18-fluorodeoxyglucose (F18-FDG) which is injected to the patient releases positrons (anti-electrons which are particles of the same properties as electrons but with positive charges). Those positrons will almost instantaneously meet an electron and have a very exo-energetic reaction (called annihilation). This annihilation transforms the whole mass of the two particles into energy and releases gamma photons in two opposite directions which are detected by the scanner cameras. The substance which is injected concentrates in the areas of the brain which are the most active, which means that those areas will exhibit a high number of annihilations. Like fMRI, the PET scan lets the neuroscientists know which areas of the brain are activated when the patient is performing an attention task. Figure 3.16 shows an example of the use of PET scan to see the influence of a flickering visual pattern in the brain.

3.2.5 Complementary Techniques to Manipulate Brain Activity: TMS or tDCS

TMS stands for transcranial magnetic stimulation and it uses electromagnetic induction to stimulate a precise region of cortex. A current passing through a coil of wire generates a magnetic field. Rapid variations of this magnetic field induce a transient electric field which in turn influences the membrane potential of nearby neurons.

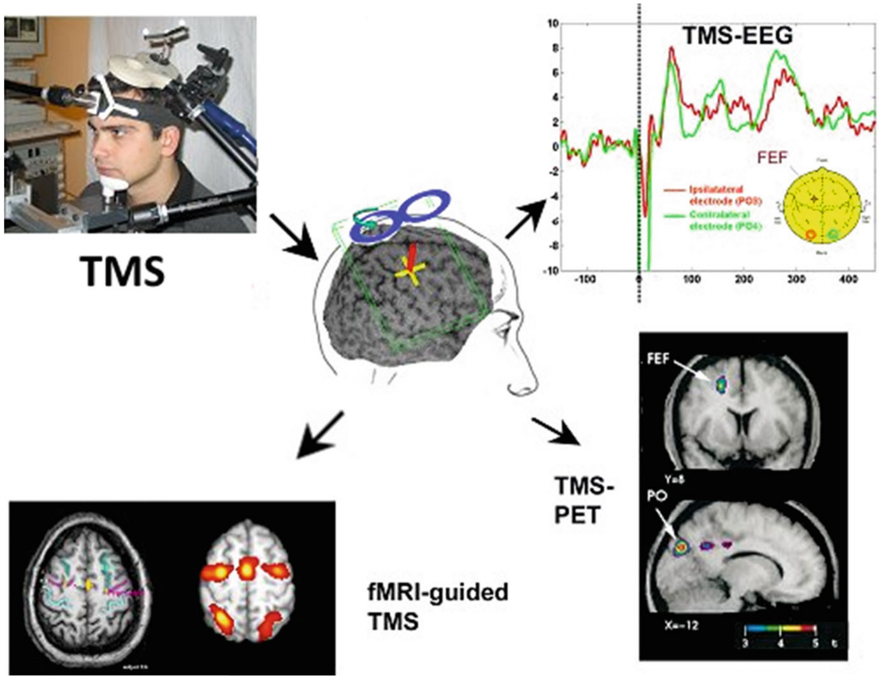

Fig. 3.17 Top left: a TMS setup. Top right: EEG modification following a TMS. Bottom left: fMRI image response after the TMS. Bottom right: PET scan response after the TMS. (Adapted from [68])

Beginning in the 1980s, TMS has been used first for clinical diagnosis and then in psychiatric therapy. It is now also used in conjunction with other imaging modalities such as fMRI and PET scans and even with EEG devices.

Imaging techniques allow us to find the active areas of the brain for a given task. However, they cannot say which part of those regions and when exactly they are really necessary to solve the task. By interfering with the normal functioning of a brain area, TMS, which has a very good spatio-temporal resolution, provides cues about when and where exactly a brain area is making its critical contribution to behavior.

Figure 3.15 shows a TMS setup application which influences EEG signals (top right), fMRI images (bottom left), and PET scan (bottom right) (Fig. 3.17).

Transcranial direct current stimulation (tDCS) is another method which aims at providing neurostimulation. The difference is that tDCS uses constant current delivered to the brain area of interest via electrodes on the scalp.

All those imaging systems let scientists manipulate attentional stimuli and measure the brain in terms of both temporal reactions and spatial localization. This provides valuable insight on attention in the brain which is very much connected to other higher-level processing such as memory or emotions and shows very complex patterns depending on the type of attention, stimuli, etc.

3.3 Summary

- Orientation of spatial attention can be overt (signaled by eye, head, and body movement) or covert (a shift of attention without over movement).
- Eye tracking remains a gold standard to detect overt attention shifts and is used in engineering and computer science, as well as in psychology and neuroscience.
- Mouse tracking is more and more used with the need to build very large stimuli datasets to model attention in computer science, especially with the arrival of deep learning.
- Covert attention shifts can be inferred from overt responses by looking at reaction time, movement trajectories, detectability, and other performance measures.
- To detect neural correlates of attention, fMRI has the best spatial resolution, and EEG/ERP and MEG the best temporal resolution.
- fMRI has become one of the most used methods in neuroscience.
- The use of TMS or tDCS in conjunction with other imaging techniques provides precise cues about when and where exactly a brain area is making its critical contribution to behavior.

References

1. Munoz, D. P. (2002). Commentary: Saccadic eye movements: Overview of neural circuitry. *Progress in Brain Research, 140*, 89–96. https://doi.org/10.1016/S0079-6123(02)40044-1. PMID: 12508584.
2. Duchowski, A. (2007). *Eye tracking methodology: Theory and practice* (Vol. 373). Springer Science & Business Media.
3. Sharafi, Z., Sharif, B., Guéhéneuc, Y. G., et al. (2020). A practical guide on conducting eye tracking studies in software engineering. *Empirical Software Engineering, 25*, 3128–3174. https://doi.org/10.1007/s10664-020-09829-4
4. Rashbass. (1960). New method for recording eye movements. *Journal of the Optical Society of America, 50*, 642–644.
5. Robinson. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions on Biomedical Engineering, 10*, 137–145.
6. Tobii eye tracking technology. http://www.tobii.com/
7. SR-Research eye tracking technology. http://www.sr-research.com/
8. Gaze Intelligence eye tracking., https://gazeintelligence.com/
9. Pupil Labs eye tracking. https://pupil-labs.com/
10. Angelopoulos, A. N., et al. (2020). Event based, near eye gaze tracking beyond 10,000 hz. *arXiv preprint arXiv:2004.03577*.
11. AdHawk mindlink eye tracking. https://www.adhawkmicrosystems.com/
12. Huang, Z., et al. (2024). Assessing the data quality of AdHawk MindLink eye-tracking glasses. *Behavior Research Methods, 56*, 1–17.

13. Hololens 2 by Microsoft. https://www.microsoft.com/en/hololens
14. Apple Vision Pro. https://www.apple.com/apple-vision-pro/
15. Huang, Z., Zhu, G., Duan, X., Wang, R., Li, Y., Zhang, S., & Wang, Z. (2024). Measuring eye-tracking accuracy and its impact on usability in apple vision pro. *arXiv preprint arXiv:2406.00255*.
16. Meta Quest Pro from Meta. https://www.meta.com/quest/quest-pro/
17. Wei, S., Bloemers, D., & Rovira, A. (2023, June). A preliminary study of the eye tracker in the meta quest pro. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences* (pp. 216–221).
18. PICO developer website, https://developer.picoxr.com/document/unreal/eye-tracking/
19. PICO website. https://www.picoxr.com
20. HTC Vive website. https://www.vive.com
21. Inseye website. https://www.inseye.com/
22. Eyetribe low cost eye-trackers. https://theeyetribe.com/
23. Gaming Tobii website. https://gaming.tobii.com/
24. Kexxu eye tracking glasses. https://kexxu.com
25. Open source recording from several eye trackers. http://www.ogama.net/
26. OpenTrack GitHub. https://github.com/opentrack/opentrack
27. GazeRecorder website. https://gazerecorder.com/
28. Beam eyewear website. https://beam.eyeware.tech/
29. WebGazer.js website. https://webgazer.cs.brown.edu/
30. Konig, S. D., & Buffalo, S. D. (2014). A nonparametric method for detecting fixations and saccades using cluster analysis: Removing the need for arbitrary thresholds. *Journal of Neuroscience Methods, 227*, 21–131.
31. Findlay, J. M., & Kapoula, Z. (1992). Scrutinization, spatial attention, and the spatial programming of saccadic eye movements. *The Quarterly Journal of Experimental Psychology A, 45*(4), 633–647.
32. Qiu, Z., et al. (2023). Fixation-related electrical potentials during a free visual search task reveal the timing of visual awareness. *Iscience, 26*, 7.
33. Poletti, M., & Rucci, M. (2015). A compact field guide to the study of microsaccades: Challenges and functions. *Vision Research Epub, 118*, 83.
34. Yuval-Greenberg, S., Merriam, E. P., & Heeger, D. J. (2014). Spontaneous microsaccades reflect shifts in covert attention. *The Journal of Neuroscience, 34*(41), 13693–13700.
35. Braun, D., & Breitmeyer, B. G. (1988). Relationship between directed visual attention and saccadic reaction times. *Experimental Brain Research, 73*(3), 546–552.
36. Kustov, A. A., & Robinson, D. L. (1996). Shared neural control of attentional shifts and eye movements. *Nature, 384*(6604), 74–77.
37. Doyle, M., & Walker, R. (2001). Curved saccade trajectories: Voluntary and reflexive saccades curve away from irrelevant distractors. *Experimental Brain Research, 139*(3), 333–344.
38. Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*(1–2), 1–37.
39. Otero-Millan, J., Castro, J. L., Macknick, S. L., & Martinez-Conde, S. (2014). Unsupervised clustering method to detect microsaccades. *Journal of Vision, 14*(2).
40. Heatmapjs, javascript API. http://www.patrick-wied.at/static/heatmapjs/
41. Simple Mouse Tracker. http://smt.speedzinemedia.com/
42. Hotjar., https://www.hotjar.com
43. Mancas, M. (2009). *Relative influence of bottom-up and top-down attention* (pp. 212–226). Attention in cognitive systems. Springer.
44. Jiang, M., et al. (2015). *SALICON: Saliency in Context*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
45. Ballard, D. H., Hayhoe, M. M., Li, F., & Whitehead, S. D. Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 337*(1281), 331–338.
46. Neuroscan website, https://compumedicsneuroscan.com/

47. Brain products website, https://www.brainproducts.com/
48. BioSemi systems website, https://www.biosemi.com/
49. NeuroSky website, https://neurosky.com/
50. Muse website, https://choosemuse.com/
51. Emotiv website, https://www.emotiv.com/
52. Müller, M. M., Gruber, T., & Keil, A. (2000). Modulation of induced gamma band activity in the human EEG by attention and visual information processing. *International Journal of Psychophysiology, 38*(3), 283–299.
53. Sauseng, P., et al. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *European Journal of Neuroscience, 22*(11), 2917–2926.
54. Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica, 42*, 313–329.
55. Sams, H., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology, 62*, 437–448.
56. Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and analysis of the component structure. *Psychophysiology, 24*, 375–425.
57. Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1991). Right hemisphere dominance of different mismatch negativities. *Electroencephalography and Clinical Neurophysiology, 78*, 466–479.
58. Paavilainen, P., Karlsson, M. L., Reinikainen, K., & Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography and Clinical Neurophysiology, 73*, 129–141.
59. Paavilainen, P., Jiang, D., Lavikainen, J., & Näätänen, R. (1993). Stimulus duration and the sensory memory trace: An event-related potential study. *Biological Psychology, 35*(2), 139–152.
60. Aaltonen, O., Niemi, P., Nyrke, T., & Tuhkahnen, J. M. (1987). Event-related brain potentials and the perception of a phonetic continuum. *Biological Psychology, 24*, 197–207.
61. Neville, H. J., & Lawson, D. (1987). Attention to central and peripheral visual space in a movement detection task: An event-related potential and behavioral study. I. Normal hearing adults. *Brain Research, 405*, 253–267.
62. Czigler, I., & Csibra, G. (1990). Event-related potentials in a visual discrimination task: Negative waves related to detection and attention. *Psychophysiology, 27*(6), 669–676.
63. Alho, K., Woods, D. L., Alagazi, A., & Näätänen, R. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology, 82*, 356–368.
64. Crottaz-Herbette, S. (2001). *Attention spatiale auditive et visuelle chez des patients héminégligents et des sujets normaux: étude clinique, comportementale et électrophysiologique*.“, PhD Thesis. University of Geneva.
65. Desmedt, J. E., & Tomberg, C. (1989). Mapping early somatosensory evoked potentials in selective attention: Critical evaluation of control conditions used for titrating by difference the cognitive P30, P40, P100 and N140. *Electroencephalography and Clinical Neurophysiology, 74*, 321–346.
66. NIMH Image library, http://infocenter.nimh.nih.gov/il/public_il/image_details.cfm?id=80
67. David Heeger courses. http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/neuroimaging/neuroimaging.html
68. Centre for Cognitive Neuroimaging (CCNi), University of Glasgow. http://www.ccni.gla.ac.uk/index.php/component/content/article/46-facilities/labs/16-ranscranial-magnetic-stimulation
69. http://www.neuroscientificallychallenged.com/glossary/functional-magnetic-resonance-imaging-fmri/