D3.3.1
Gaze Interaction and Analytics

Work package	**WP3 Multimodal Analytics and Assessment**
Task	**T3.3 Gaze Interaction and Analytics**
Editor	**Lucas Paletta (JRD)**
(co-)authors	**TBD**
Public / confidential	**public**

Project PLAYTIME

The research leading to these results has received funding from the AAL Programme of the European Union and by the Austrian BMVIT/FFG under the Agreement no 857334, ZonMw (the Netherlands) and funding from Brussels (Belgium). It reflects only the author's view and that the Union is not liable for any use that may be made of the information contained therein.

28/02/2018

PLAYTIME partner			organisation
01	JR		JOANNEUM RESEARCH Forschungsgesellschaft mbH DIGITAL – Institut für Informations- und Kommunikationstechnologien, 8010 Graz
02	FAM		FameL GmbH Steinbruchweg 20, A-8054 Seiersberg
03	SVD		Sozialverein Deutschlandsberg Kirchengasse 7, A-8543 Deutschlandsberg
04	GGZ		Geestelijke Gezondheidszorg Eindhoven en de Kempen Postbus 909, 5600 AX Eindhoven, The Netherlands
05	TIU		Stichting Katholieke Universiteit Brabant, Tilburg University PO Box 90153, 5000 LE Tilburg, The Netherlands
06	MCR		McRoberts BV. Raamweg 43, 2596 HN The Hague, The Netherlands
07	MBY		MindBytes F. Roosseveltlaan 348-349, B8, 90600 Ghent, Belgium
08	GEU		Ghent University Sint-Pietersnieuwstraat 25, 9000 Gent, Belgium

Contents

1 Executive Summary

This is the first of a set of three PLAYTIME Deliverables D3.3.1-D3.3.3 on the Task T3.3 Gaze Interaction & Analytics.

The Deliverable provides an introduction into eye tracking technologies, motivates the use of technology that is applied in PLAYTIME, and provides insight into the background, objectives and implementation of gaze analytics as well as gaze interfaces for attention games (MIRA).

2 Introduction

Eye tracking provides a useful methodology for monitoring changes in cognition and giving fundamental insight into human behavior (Kaufmann et al., 2012) by detecting executive deficits detected in Alzheimer's disease for example using the anti-saccade task.

Crawford et al. (2005) and others suggest that this approach has a number of advantages over traditional methods of psychological assessment. Standard and novel experimental paradigms can be developed to evaluate theories of cognitive impairment and to dissociate between various neural networks.

Objective of this task performed by JRD is the provision of real-time and on-line available analytics about relevant diagnostic features gained from eye movements and dependent behavior analysis. Features about characteristics of interaction are captured in the context of visual cueing, executive functions and sensorimotor performance as well, taking the psychological framework of Task 3.1 into account.

3 Attention in the context of dementia

3.1 Neuropsychological findings on attention and Alzheimer

In the critical review by (Perry & Hodges, 1999), the progress that has been made in the research on attentional and executive deficits in Alzheimer's disease has been summarized. Like memory, attention was recognized is consisting of subtypes that differ in their function and anatomical basis. The review bases the review upon a classification of three subtypes of attention: selective, sustained and divided. This model is derived from lesion studies, animal electrophysiological recordings and functional imaging. In the review is examined how these subcomponents of attention can be reconciled with neuropsychological models of attentional control, particularly the Supervisory Attentional System and the Central Executive System of Shallice and Baddeley, respectively: Substantial evidence suggests that after an initial amnesic stage in Alzheimer's disease, attention is the first non-memory domain to be affected, before deficits in language and visuospatial functions. They comment that this would be consistent with the possibility that difficulties with activities of daily living, which occur in even mildly demented patients, may be related to attentional deficits. It appears that divided attention and aspects of selective attention, such as set-shifting and response selection, are particularly vulnerable while sustained attention is relatively preserved in the early stages.

The phenomenon of cognitive slowing in Alzheimer's disease and normal ageing emphasizes the need to discriminate quantitative changes in attention dysfunction from qualitative changes which may be specifically related to the disease process. The neuropathological basis of these attentional deficits remains unsettled, with two competing hypotheses: spread of pathology from the medial temporal to basal forebrain structures versus corticocortical tract disconnection. They finally discussed the difficulties of comparing evidence across studies and looked at the implications for the design of future studies and future directions that might be fruitful in the research on attention in Alzheimer's disease.

3.2 Dementia and executive functions

Dementia with its predominate subtypes, i.e., Alzheimer’s disease (AD) and vascular dementia (VaD), could lead to progressive decline in executive function as well as other cognitive functions. The impact of these symptoms on the life of patients and caregivers are devastating, requiring an adequate planning ahead for long-term management issues (Denckla, 1994). Executive dysfunction may be an early feature of some AD patients (McPherson & Simpson, 2002). The structural correlate to certain dimensions of executive dysfunction in AD patients has been suggested to relate to changes in the deep frontal white matter (Sjöbeck et al., 2010). VaD is often implicated in executive dysfunction such as progressive subcortical vascular encephalopathy and selective white matter hyper-intensities (WMHs). Executive dysfunction significantly predicts future decline in this dementia (Boyle & Cahn-Weiner, 2005).

3.2.1 Executive functions

Executive functions (EFs; also called executive control or cognitive control) refer to a family of top-down mental processes needed when you have to concentrate and pay attention, when going on automatic or relying on instinct or intuition would be ill-advised, insufficient, or impossible (Burgess & Simons 2005, Espy 2004, Miller & Cohen 2001). Using EFs is effortful; it is easier to continue doing what you have been doing than to change, it is easier to give into temptation than to resist it, and it is easier to go on “automatic pilot” than to consider what to do next. There is general agreement that there are three core EFs (e.g., Lehto et al. 2003, Miyake et al. 2000):

- inhibition [inhibitory control, including self-control (behavioral inhibition) and interference control (selective attention and cognitive inhibition)],
- working memory (WM), and
- cognitive flexibility (also called set shifting, mental flexibility, or mental set shifting and closely linked to creativity).

From these, higher order EFs are built such as reasoning, problem solving, and planning (Collins & Koechlin 2012, Lunt et al. 2012). EFs are skills essential for mental and physical health; success in school and in elderly life; and cognitive, social, and psychological development.

3.2.2 Inhibitory control

Inhibitory control (one of the core EFs) involves being able to control one’s attention, behavior, thoughts, and/or emotions to override a strong internal predisposition or external lure, and instead do what’s more appropriate or needed. Without inhibitory control we would be at the mercy of impulses, old habits of thought or action (conditioned responses), and/or stimuli in the environment that pull us this way or that. Thus, inhibitory control makes it possible for us to change and for us to choose how we react and how we behave rather than being unthinking creatures of habit. It doesn’t make it easy. Indeed, we usually are creatures of habit and our behavior is under the control of environmental stimuli far more than we usually realize, but having the ability to exercise inhibitory control creates the possibility of change and choice. It can also save us from making fools of ourselves.

Inhibitory control of attention (interference control at the level of perception) enables us to selectively attend, focusing on what we choose and suppressing attention to other stimuli. We need such selective attention at a cocktail party when we want to screen out all but one voice. A salient stimulus such as visual motion or a loud noise attracts our attention whether we want it to or not. That is called exogenous, bottom-up, automatic, stimulus-driven, or involuntary attention and is driven by properties of stimuli themselves (Posner & DiGirolamo 1998, Theeuwes 1991). We can also choose voluntarily to ignore (or inhibit attention to) particular stimuli and attend to others based on our goal or intention. Besides being called selective or focused attention, this has been termed attentional control or attentional inhibition, endogenous, top-down, active, goal-driven, voluntary, volitional, or executive attention (Posner & DiGirolamo 1998, Theeuwes 2010).

Another aspect of interference control is suppressing pre-potent mental representations (cognitive inhibition). This involves resisting extraneous or unwanted thoughts or memories, including intentional forgetting (Anderson & Levy 2009), resisting proactive interference from information acquired earlier (Postle et al. 2004), and resisting retroactive interference from items presented later. Cognitive inhibition is usually in the service of aiding WM and is discussed in the section Inhibitory Control Supports Working Memory. It tends to cohere more with WM measures than with measures of other types of inhibition.

Self-control is the aspect of inhibitory control that involves control over one's behavior and control over one's emotions in the service of controlling one's behavior. Self-control is about resisting temptations and not acting impulsively. The temptation resisted might be to indulge in pleasures when one should not (e.g., to indulge in a romantic fling if you are married or to eat sweets if you are trying to lose weight), to overindulge, or to stray from the straight and narrow (e.g., to cheat or steal). Or the temptation might be to impulsively react (e.g., reflexively striking back at someone who has hurt your feelings) or to do or take what you want without regard for social norms (e.g., butting in line or grabbing another child's toy).

Another aspect of self-control is having the discipline to stay on task despite distractions and completing a task despite temptations to give up, to move on to more interesting work, or to have a good time instead. This involves making yourself do something or keep at some- thing though you would rather be doing some- thing else. It is related to the final aspect of self-control — delaying gratification (Mischel et al.1989) — making yourself forgo an immediate pleasure for a greater reward later (often termed delay discounting by neuroscientists and learning theorists; Louie & Glimcher 2010, Rachlin et al. 1991). Without the discipline to complete what one started and delay gratification, no one would ever complete a long, time-consuming task such as writing a dissertation, running a marathon, or starting a new business.

Although the above examples typically involve a tug-of-war between a part of you that wants to do x and another part of you that wants to do y (Hofmann et al. 2009), self-control can be needed where there are not competing desires. It is needed, for example, to not blurt out what first comes to mind (which might be hurtful to others or embarrassing to you), to not jump to a conclusion before getting all the facts, or to not give the first answer that occurs to you when if you took more time you could give a better, wiser response.

Errors of impulsivity are errors of not being able to wait. If someone can be helped to wait such errors can often be avoided. Many of us have had the experience of pressing the "send" button for an email only to wish we had not. Many of us have also had the experience of our first interpretation of the intention behind someone's words or actions being incorrect, and we have either been grateful we exercised the self-control to wait until we acquired more information or regretted that we acted precipitously without waiting. On laboratory tasks, young children often rush to respond and thus make errors by giving the pre-potent response when a different response is required. Helping young children wait improves their performance. This has been shown using a variety of inhibitory control tasks such as go/no-go (Jones et al. 2003), theory of mind (Heberle et al., 1999), day-night (Diamond et al. 2002), and a Piagetian search task (Riviere & Lecuyer, 2003). The sub-thalamic nucleus appears to play a critical role in preventing such impulsive or premature responding (Frank 2006).

Examples of typical psychological measures of inhibitory control include the Stroop task (MacLeod 1991), Simon task (Hommel 2011), Flanker task (Eriksen & Eriksen 1974, Mullane et al. 2009), antisaccadic tasks (Luna 2009, Munoz & Everling 2004), delay-of-gratification tasks (Kochanska et al. 2001, Sethi et al. 2000), go/no-go tasks (Cragg & Nation 2008), and stop-signal tasks (Verbruggen & Logan 2008). One of the many hotly debated aspects of EFs is which component(s) of EFs a task requires.

3.2.3 Eye tracking and dementia

Progressive neurological diseases, such as, the Alzheimer, Parkinson, Huntington or Wilson, are well known for the decrease in eye movement behavior (Kuskowski, 1988; White et al., 1983). The characteristics of the impairment support clinicians to localize brain lesions as well as to determine diagnostics about the trajectory of the diseases (Kuskowski, 1988). Dysfunctionality in the continuous tracking of stimuli was already associated with Alzheimer dementia by (Fletcher, 1988).

(Crawford et al., 2005) has identified the important indication that Alzheimer patients are characterized with a significant impairment of their inhibitory functionality of eye movements, due to neurodegeneration of frontal and prefrontal lobes which are responsible for inhibitory effects (Kaufmann et al., 2010; Pierrot-Deseilligny et al., 2004).

In early stages of Alzheimer disease, the antisaccade task is known to identify Alzheimer. This task requires from the test person a voluntary turning away from an actual stimulus and analyses the eye movement behavior further (Kaufman, et al., 2010).

Another recent study (Crutcher et al., 2009) showed the promise of the Visual Paired Comparison (VPC) task for the detection of memory impairment associated with MCI. There are two phases for each trial of the VPC task. First, during the familiarization phase, subjects are presented with two identical visual stimuli, side by side, on a computer screen. Subjects are allowed to look at the pictures for a specified amount of time. During the test phase, which follows a variable delay, subjects are presented with pictures of the old stimulus and a novel stimulus, side by side. Eye movements are monitored via noninvasive infrared eye tracking, and control subjects typically spend 70% of the time during the test phase looking at the novel stimulus. This indicates that they have a memory for the repeated, and now less interesting, stimulus. In contrast, it has been shown for age-matched MCI patients in (Lagun et al., 2011) that they did not spend more time looking at the novel stimulus than the repeated stimulus (Crutcher et al., 2009), suggesting they did not remember which stimulus was novel and which was familiar.

(Lagun et al., 2011) applied automatic machine learning methods from computer science to analyze and exploit the information contained in the characteristics of eye movement exhibited by healthy and impaired subjects during the viewing of stimuli in the VPC task. Specifically, they hypothesized that additional characteristics of eye movement would help improve classification accuracy of cognitive impairment, thus allowing classification algorithms to more accurately distinguish healthy from impaired subjects. They first trained the classification models on the multidimensional representation of eye movements from a sample of the impaired and control subjects, and then used the model to predict the status of new subjects based on their eye movement characteristics. The results show that eye movement characteristics including fixation duration, saccade length and direction, and re-fixation patterns (defined in next section) can be used to automatically distinguish impaired and normal subjects. Accordingly, this generalized approach may be useful for improving early detection of AD, and may be applied, in combination with other behavioral tasks, to examine cognitive impairments associated with other neurodegenerative diseases.

3.3 Attention, goal striving and motivation

3.3.1 Motivational reserve and dementia

Motivation is an umbrella term for various processes involved in goal-directed behaviour (Braver et al., 2014). It has been suggested by early personality psychologists (Lewin et al., 1944) and more differentiated in current models of motivation (Heckhausen & Heckhausen, 2008) that two main motivational phases can be distinguished: goal setting and goal striving. Goal setting and striving are determined by rather different motivation-related constructs (Gollwitzer et al., 2012).

While goal setting is determined mainly by control and expectancy constructs (Skinner, 1996) such as self-efficacy (Bandura, 1997), goal striving is rather determined by volitional or self-regulatory strategies that are needed to cope with difficulties during the implementation phase such as decision regulation (Kuhl & Fuhrmann, 1998), activation regulation (Kruglanski et al., 2000), and motivation regulation (Kuhl & Fuhrmann, 1998). Further self-regulatory strategies are important during goal striving, e.g., emotion and **attention regulation**. In total, there are four basic motivation-related processes relevant in goal setting and striving that are usually measured by self-report, and further variables relating to these four processes, but measured by behavioral testing, scenario tests, or an occupation-based scoring procedure.

Studies on the neuropathology of Alzheimer's disease (AD) have repeatedly shown that many individuals with pronounced neuropathological, AD-typical changes in the brain exhibit no clinical manifestation of dementia syndrome (Knopman et al., 2003). One explanation for this finding is the concept of brain reserve (Fratiglioni et al., 2007; Stern, 2006; Valenzuela & Sachdev, 2006). Brain reserve can be defined as the ability of the brain to tolerate or compensate for age- and disease-related changes in a way that cognitive function is still maintained. In order to clinically manifest impairments in the cognitive and functional abilities of an individual with a larger brain reserve, it is assumed that the neuropathological damage must be more severe. The neuropathological processes seem to accumulate until they are severe enough to cross a threshold and be reflected in the clinical picture (Stern, 2006). Concepts of brain reserve can explain why people differ in their reserve capacity. The extension of this threshold of clinical manifestation and the increase in brain reserve is one of the potential preventive goals for AD (Kivipelto & Solomon, 2008).

The model of (Forstmeier & Maercker, 2015) presumes that **activating motivational processes** during the life course rises the number of synaptic connections and stimulates the development of new neurons. These neurophysiological alterations increase the efficacy of usage of relevant brain networks and enable the brain to compensate for disrupted networks (this is captured in the term 'brain reserve'). There is a wealth of evidence that the human brain still exhibits plasticity in adult and older life (Kempermann et al., 2002). The brain areas primarily involved in motivational processes are the amygdala (fear-motivated behavior), the nucleus accumbens (reward-motivated behavior), and the prefrontal cortex (regulating motivational salience and determining intensity of responding) (Cardinal et al., 2002; Kalivas & Volkow, 2005).

Correspondingly, it was used the term motivational reserve (MR) to describe the effect of motivational activities on reserve capacity (Maercker & Forstmeier, 2011). MR can be defined as a set of motivational abilities or processes that provide the individual with resilience to neuropathological deterioration. Motivational and cognitive reserve constitute complementary concepts in (Forstmeier & Maercker, 2015)'s model. Their model of includes additional factors that might mediate the effect of MR on further brain areas. These factors act by influencing stress activation, vascular risk factors, cognitive training, and **emotional health** (Forstmeier & Maercker, 2008; Forstmeier et al., 2012).

There is evidence of an association between depressive symptoms and a higher rate of conversion form MCI to dementia (Gabryelewicz et al., 2007; Modrego & Ferrandez, 2004). However, there is evidence that **motivation-related symptoms** of depression have a higher predictive value than the affect-related symptoms (Berger et al., 2002; Bartolini et al., 2005). Therefore, it is not surprising that apathy is also associated with a higher rate of conversion form MCI to dementia (Robert et al., 2006; Teng et al., 2007).

Enhancing motivational abilities might also increase the effectiveness of pharmacological and psychosocial interventions. By targeting the neuro- psychiatric symptoms of apathy and depression, care- giver burden might be reduced. The results of previous studies on brain reserve have been used to make recommendations about lifestyle and activities in late life to decrease the risk of AD (Fratiglioni et al., 2004). The possible results the model and study of (Forstmeier & Maercker, 2015) would add recommendations about changes in motivational self-regulation in later life (Maercker & Forstmeier, 2011).

3.3.2 Motivational resources and attention

The findings in (Engelman et al., 2009) support the notion that **motivation improves behavioral performance in a demanding attention task** by enhancing evoked responses across a distributed set of anatomical sites, many of which have been previously implicated in attentional processing. (Engelman et al., 2009) represents the first study that focuses on and proves motivational variables as predictors of cognitive decline and emotional health in the cognitively impaired. Therefore our work will focus on the impact of motivational variables, like measurement of engagement of the person with dementia, in the relation with the performance of attentional processes, such as, described in more detail in Chapter 6.

4 Measurement of attention

4.1 Eye tracking technologies

4.1.1 Eye movement analysis

Eye tracking refers to the process of tracking eye movements or the absolute point of gaze (POG) — referring to the point the user's gaze is focused at in the visual scene. Eye tracking is useful in a broad range of application areas, from psychological research and medical diagnostic to usability studies and interactive, gaze-controlled applications. This chapter is focused on the use of real-time data from human eye movements that can be exploited in the area of human-technology interaction. The aim is to provide a useful starting point for researchers, designers of interactive technology and assistive technology professionals wishing to gain deeper insight into gaze interaction.

Initially, eye movements were mainly studied by physiological introspection and observation. Basic eye movements were categorized and their duration estimated long before the eye tracking technology enabled precise measurement of eye movements. The first generation of eye tracking devices was highly invasive and uncomfortable. A breakthrough in eye tracking technology was the development of the first ‘‘non-invasive’’ eye tracking apparatus in the early 1900s (Wade and Tatler, 2005), based on photography and light reflected from the cornea. It can be considered as the first ancestor of the current widely used video-based, corneal reflection eye tracking systems. The development of unobtrusive camera-based systems (Morimoto and Mimica 2005) and the increase of computing power enabled gathering of eye tracking data in real time, enabling the use of gaze as a control method for people with disabilities (Ten Kate et al. 1979; Friedman et al. 1982). Since then, eye tracking has been used in a wide range of application areas, some of which are reviewed later in this chapter.

Using the eye as an input method has benefits but also some considerable challenges. These challenges originate from eye physiology and from its perceptive nature. Below, we briefly introduce the basics of eye physiology and eye movements. The rest of the chapter concentrates on giving an overview of eye tracking technology and methods used to implement gaze interaction. We will also review related research and introduce example applications that should help the readers to understand the reasons behind the problem issues—and to design solutions that avoid the typical pitfalls. We will conclude this chapter with a summary and a discussion of potential future research directions for gaze-based interfaces.

To see an object in the real world, we have to fixate our gaze at it long enough for the brain's visual system to perceive it. Fixations are often defined as pauses of at least 100 ms, typically between 200 and 600 ms. During any one fixation, we only see a fairly narrow area of the visual scene with high acuity. To perceive the visual scene accurately, we need to constantly scan it with rapid eye movement, so-called saccades. Saccades are quick, ballistic jumps of 2° or longer that take about 30–120 ms each (Jacob 1995). In addition to saccadic movements, the eyes can smoothly follow a moving target; this is known as (smooth) pursuit movement. For more information about other types of eye movements and visual perception in general, see, for example, Mulvey (2012).

The size of the high-acuity field of vision, the fovea, subtends at an angle of about one degree from the eye. The diameter of this region corresponds to an area of about two degrees, which is about the size of a thumbnail when viewed with the arm extended (Duchowski and Vertegaal 2000). Everything inside the fovea can be perceived with high acuity but the acuity decreases rapidly towards the periphery of the eye. The cause for this can be seen by examining the physiology of the retina (see Figure 1). The lens focuses the light coming from the pupil on the center of the retina. The fovea is packed with photoreceptive cells but the density of these cells decreases rapidly in the peripheral area. The fovea mainly contains cones, photoreceptive cells that are sensitive to color and provide acuity. In contrast, the peripheral area contains mostly rods, i.e. cells that are sensitive to light, shade and motion. The remaining peripheral vision provides cues about where to look next and also gives information on movement or changes in the scene in front of the viewer. For example, a sudden movement in the periphery can thus quickly attract the viewer's attention (Hillstrom and Yantis 1994).

We only see a small fraction of the visual scene in front of us with high acuity at any point in time. The need to move our eyes toward the target is the basis for eye tracking: it is possible to deduct the gaze vector by observing the ‘‘line-of-sight’’

4.1.2 Eye Tracking Techniques

While a large number of different techniques to track eye movements have been investigated in the past, three eye tracking techniques have emerged as the pre- dominant ones and are widely used in research and commercial applications today. These techniques are (1) videooculography (VOG), video based tracking using head-mounted or remote visible light video cameras, (2) video-based infrared (IR) pupil-corneal reflection (PCR), and (3) Electrooculography (EOG). While particularly the first two video-based techniques have a lot of properties in common, all techniques have application areas where they are most useful.

Video-based eye tracking relies on off-the-shelf components and video cameras and can therefore be used for developing ‘‘eye-aware’’ or attentive user interfaces that do not strictly require accurate point of gaze tracking (e.g. about 4°, Hansen and Pece 2005). In contrast, due to the additional information gained from the IR-induced corneal reflection, IR-PCR provides highly accurate point of gaze measurements of up to 0.5° of visual angle and has therefore emerged as the preferred technique for scientific domains, such as usability studies or gaze-based interaction, and commercial applications, such as in marketing research. Finally, EOG has been used for decades for ophthalmological studies as it allows for measuring relative movements of the eyes with high temporal accuracy. In addition to different application areas, each of these measurement techniques also has specific technical advantages and disadvantages that we will discuss in the following sections.

4.1.3 Video based tracking

A video-based eye tracking system can be either used in a remote or head-mounted configuration. A typical setup consists of a video camera that records the movements of the eye(s) and a computer that saves and analyses the gaze data. In remote systems, the camera is typically based below the computer screen (Figure 1) while in head-mounted systems, the camera is attached either on a frame of eyeglasses or in a separate ‘‘helmet’’. Head-mounted systems often also include a scene camera for recording the user’s point of view, which can then be used to map the user’s gaze to the current visual scene.

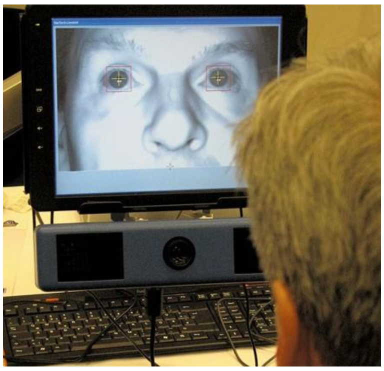


Figure 1. The eye tracker’s camera is placed under the monitor. Infrared light sources are located on each side of the camera. IR is used to illuminate the eye and its reflection on the cornea provides an additional reference point that improves accuracy when tracked together with the center of the pupil (© 2008, www.cogain. org)

The frame rate and resolution of the video camera have a significant effect on the accuracy of tracking; a low-cost web camera cannot compete with a high-end camera with high-resolution and high sample rate. The focal length of the lens, the angle, as well as the distance between the eye and the camera have an effect on the working distance and the quality of gaze tracking. With large zooming (large focal length), it is possible to get a close-up view of the eye but it narrows the working angle of the camera and requires the user to sit fairly still (unless the camera follows the user’s movements). In head-mounted systems, the camera is placed near the eye, which means a bigger image of the eye and thus more pixels for tracking the eye. If a wide angle camera is used, it allows more freedom of movement of the user but also requires a high-resolution camera to maintain enough accuracy for tracking the pupil (Hansen and Majaranta 2012).

Since tracking is based on video images of the eye, it requires an unobstructed view of the eye. There are a number of issues that can affect the quality of tracking, such as varying light conditions, reflections of eyeglasses, droopy eyelids, squinting the eyes while smiling, or even heavy makeup (for more information and guidelines, see Goldberg and Wichansky 2003).

The video images are the basis for estimating the gaze position on the computer screen: the location of the eye(s) and the center of the pupil are detected. Changes in their position are tracked, analyzed and mapped to gaze coordinates. For a detailed survey of video-based techniques for eye and pupil detection and gaze position estimation, see Hansen and Ji (2009). If only the pupil center is used and no other reference points are available, the user must stay absolutely still for an accurate calculation of the gaze vector (the line of sight from the user’s eye to the point of view on the screen.).

4.1.4 Infrared Pupil-Corneal Reflection Tracking

Systems only based on visible light and pupil center tracking tend to be inaccurate and sensitive to head movement. To address this problem, a reference point, a so called ‘‘corneal reflection’’ or glint, can be added. Such a reference point can be added by using an artificial infrared (IR) light source aimed on- or off-axis at the eye. An on-axis light source will result in a ‘‘bright pupil’’ effect, making it easier for the analysis software to recognize the pupil in the image. The effect is similar to the red-eye effect caused by flash in a photograph. The off-axis light results in ‘‘dark pupil’’ images. Both will help in keeping the eye area well lit but they do not disturb viewing or affect pupil dilation since IR light is invisible to the human eye (Duchowski 2003).

By measuring the corneal reflection(s) from the IR source relative to the center of the pupil, the system can compensate for inaccuracies and also allow for a limited degree of head movement. Gaze direction is then calculated by measuring the changing relationship between the moving pupil center of the eye and the corneal reflection (see Fig. 3.3). As the position of the corneal reflection remains roughly constant during eye movement, the reflection will remain static during rotation of the eye and changes in gaze direction, thus giving a basic eye and head position reference. In addition, it also provides a simple reference point to compare with the moving pupil, and thus enables calculation of the gaze vector (for a more detailed explanation, see Duchowski and Vertegaal 2000).

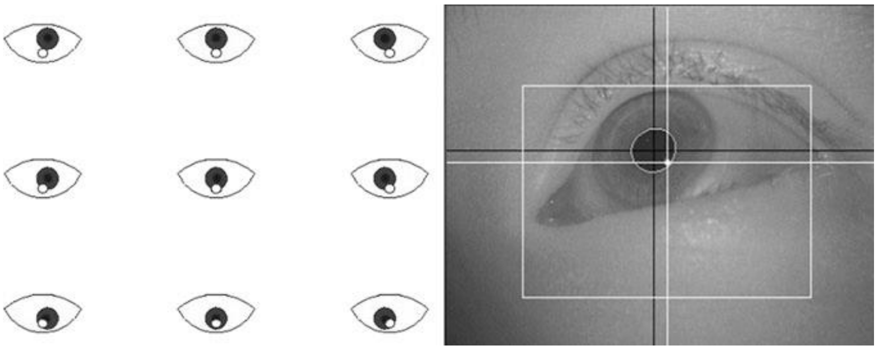


Figure 2. The relationship between the pupil center and the corneal reflection when the user fixates on different locations on the screen (Adapted from Majaranta et al. 2009b)

While IR illumination enables fairly accurate remote tracking of the user it does not work well in changing ambient light, such as in outdoors settings. There is an ongoing research that tries to solve this issue (see e.g. Kinsman and Pelz 2012; Bengoechea et al. 2012). In addition, according to our personal experience, there seems to be a small number of people for whom, robust/accurate eye tracking does not seem to work even in laboratory settings. Electrooculography is not dependent or disturbed by lighting conditions and thus can replace VOG-based tracking in some of these situations and for some applications.

4.1.5 Eye Tracker Calibration and Accuracy

Before a remote video-based eye tracking system can map gaze onto a screen, it must be calibrated to that screen for each user. This is usually done by showing a number of calibration points on the screen and asking the user to consecutively fixate at these points, one at a time.

The relationship between the pupil position and the corneal reflections changes as a function of eye gaze direction (see Figure 4). The images of the eye and thus its orientation in space are analyzed by the computer for each calibration point, and each image is associated with corresponding screen coordinates. These main points are used to calculate any other point on-screen via interpolation of the data.

Calibration is a key factor defining the accuracy of any eye tracker. If the calibration is successful, the accuracy is typically about 0.5° of visual angle. This corresponds to about 15 pixels on a 1700 display with the resolution of 1024 9 768 pixels, viewed from a distance of 70 cm. Even with successful calibration, the practical accuracy may be less due to drifting which will cause an offset between the measured point of gaze and the actual gaze point. Such drifting may be caused by changes in lighting and pupil size. In head mounted systems, it is also possible that the camera has moved along with the frames. Various methods have been implemented to prevent drifting (Hansen et al. 2010) or to cope with it (Stampe and Reingold 1995). If both eyes are tracked, it usually not only improves accuracy in general but also limits drifting. Since calibration takes time and may be seen as an obstacle for using eye tracking in everyday applications, techniques requiring only one calibration point, automatic calibration procedures, and ‘‘calibration-free’’ systems have been developed (see e.g. Nagamatsu et al. 2010).

Even if the eye tracker was perfectly accurate, it may still be impossible to know the exact pixel the user is focused on. This is because everything within the fovea is seen in detail and the user can move attention within this area without voluntarily moving her eyes. Besides, the eyes perform very small, rapid movements, so-called micro saccades, even during fixations to keep the nerve cells in the retina active and to correct slight drifting in focus. Thus, if the cursor of an ‘‘eye mouse’’ were to follow eye movements faithfully, the cursor movement would appear jerky and it would be difficult to concentrate on pointing (Jacob 1993). Therefore, the coordinates reported by the system are often ‘‘smoothed’’ by averaging data from several raw gaze points. This may have an effect on the responsiveness of the system, especially if the sample rate of the camera is low.

4.1.6 Choice of eye tracking technology

Eye tracking is no longer a niche technology used by specialized research laboratories or a few select user groups but actively exploited in a wide variety of disciplines and application areas. When choosing an eye tracking system, one should pay attention to the hardware’s gaze tracking features as well as the accompanying software and additional accessories. Many eye tracking manufacturers provide different models targeted at different purposes. The systems may use the same basic technical principles of operation, but what makes a certain system suitable for a specific purpose are the applications (software) that utilize the raw eye data, e.g. software for recording and analyzing the gaze path, or assistive software that allow the eye to be used as an substitute for the mouse.

Issues to consider from the technical part that affect the suitability of the system for a specific purpose include: spatial and temporal resolution (accuracy), camera angle, freedom of head movements, tolerance to ambient light, tolerance to eye glasses and contact lenses, possibility to track only one or two eyes.

Video-based eye tracking, especially if implemented as a remote tracker, provides a fairly comfortable non-invasive (contact-free) option for the users. Systems that combine the video with infrared (i.e. track both the pupil and the IR corneal reflection) also provide reasonable freedom of head movement without sacrificing the accuracy too much. However, those systems, especially IR-PCR, are very sensitive to ambient light and changes in the light levels and only provide limited temporal accuracy and recording time. EOG-based eye trackers are not sensitive to lighting conditions. The downside is that EOG can be considered invasive and may be seen as impractical for everyday use, because it require electrodes to be placed around the eye to measure the skin's electrical potential differences.

There are also differences in the data produced by each of the trackers. Since an EOG-based system provides information on relative eye movements, it is especially useful in situations where only changes in the gaze direction are required (e.g. gaze gestures, navigation and steering tasks, or research on saccades and smooth pursuit). However, VOG based systems may work better if an exact point of gaze is important (e.g. point-and-select tasks).

For some tasks, a combination of EOG and VOG might provide the best results.

Apart from a few exceptions (e.g. Du et al. 2012) their combined use has not been applied much. It has been recommended in Bulling et al., 2012 that there might be high potential in using a method that combines the best features of each technology, especially for passive eye monitoring and clinical studies conducted in challenging outdoor environments.

From the human factors point of view, the system's invasiveness, ease of use, setup time and available customer support are important issues. For a disabled person, an eye control system is a way of communicating and interacting with the world and may be used extensively in varying conditions. Thus, reliability, robustness, safety, and mounting issues must be carefully taken into account, in addition to ease of use and general usability. In addition, one should be able to tailor the system to match each user's abilities, needs and challenges induced by disease (Donegan et al. 2009). With increased availability, reliability and usability of the state-of-the-art trackers, the focus on gaze assistive technology research is slowly moving from technical challenges toward the human point of view, presenting a need to also study user experience (Mele and Federici 2012).

Eye tracking is becoming an increasingly interesting option even in traditional computing. Major technology companies and the gaming industry are starting to show growing interest in embedding eye tracking in their future products, such as laptops and tablets (Tobii 2011; Fujitsu 2012). Vision based technologies are already widely used in the gaming field, enabling players to use gestures and full body movement to control the games, and eye tracking is envisioned to be part of future gaming (Larsen 2012). The hype on mixed reality and smart glasses (such as the HoloLens, together with eye tracking devices) indicates that it is only a matter of time, when the first broader wave of eye-controlled consumer product will enter the market.

4.1.7 Gaze Interaction Applications

Information from eye movements and gaze direction can be utilized on various levels in a wide variety of applications. Some applications require the user to move her eyes voluntarily while other systems monitor the eye movements subtly in the background and require no special effort from the user.

Fairclough (2011) suggested a four-group classification to describe the different kinds of physiological computing system. The categories start from the overt, intentional systems (such as conventional input devices) and end with the covert, unintentional systems (e.g. ambulatory monitoring). Each individual system can be placed on this continuum and some systems may also be hybrids that contain features from different categories. A similar continuum can also be used to describe different types of gaze-based systems, starting from applications where the intention is the driving force, requiring full, overt, conscious attention from the user (see Fig. 3.6). In the middle, we have attentive applications that somehow react to the user's behavior but do not require any explicit control from the user. An advanced form of this are adaptive applications that learn the user's behavior

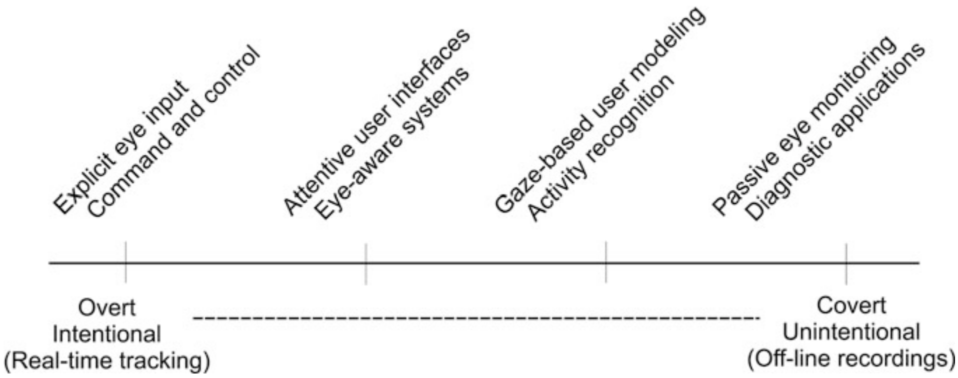


Figure 3. Continuum of eye tracking applications from intentional to unintentional systems

patterns and are able to model the user's behavior. In the other end, we have systems that passively monitor their eye behavior, requiring no conscious input from the user whose behavior is monitored and analyzed. In the following section, we provide examples of representative applications for each category.

(1) Explicit eye input is utilized in applications that implement gaze-based command and control. Here, people use voluntary eye movements and consciously control their gaze direction, for example, to communicate or control a computer. In its simplest form, the eye can be used as a switch. For example, the user may blink once or twice, or use simple vertical or horizontal eye movements as an indication of agreement or disagreement (obtainable even with a low-cost web camera based tracker). The most common way to implement gaze-based control is to use the eye's capability to point at the desired target (requiring a more accurate tracker). Mouse emulation, combined with different techniques and gaze-friendly tools for dragging, double- click, screen magnification etc. make it possible to control practically any graphical interface based on windows, icons, menus, and pointer devices (WIMP). Example applications include gaze-based text entry, web browsing, gaze- controlled games, music, etc. (For a review of different techniques and applications for gaze-based computer control, see (Skovsgaard et al. 2012)).

(2) Attentive user interfaces can be considered as non-command interfaces (Nielsen 1993) where the user is not expected to change his or her gaze behavior to give explicit commands. Instead, the information of the user's natural eye movements is used subtly in the background. In its simplest form, an attentive interface may implement a ‘‘gaze-contingent display’’ that shows a higher resolution image on the area the user is focusing

on while maintaining a lower resolution in the periphery in order to save bandwidth (Duchowski et al. 2004). As the application reacts to the user's natural behavior by explicit commands, it is important to provide enough information of the system state so that the user does not lose control and is able to react to potential problems caused by the inaccuracy of gaze. It has been argued (e.g. by Jacob 1993), that the eye, as a perceptual organ, is best suited for interaction as an additional input. For a review and more information about attentive applications, see Räihä et al. (2011) or Istance and Hyrskykari (2012).

(3) Gaze-based user modeling provides a way to better understand the user's behavior, cognitive processes, and intentions. All of the previous gaze-based applications explicitly or implicitly assume that the sole entity of interest is the user's point of gaze on a specific interactive surface or interface. In addition, the vast majority of these applications use gaze as an explicit input. Automated analysis of eye movements has a long history as a tool in experimental psychology to better understand the underlying cognitive processes of visual perception. For example, Markov processes have been used to model fixation sequences of observers looking at objects with the goal of quantifying the similarity of eye movements (Hacisalihzade et al. 1992), to identify salient image features that affected the perception of visual realism (ElHelw et al. 2008), or to interpret eye movements as accurately as human experts but in significantly less time (Salvucci and Anderson 2001). A recent trend in human attention analysis investigates a small set of specific eye movement characteristics toward developing holistic computational models of a user's visual behavior. These models typically rely on computational methods from machine learning and pattern recognition (Kandemir and Kaski 2012). As evidenced by research in experimental psychology, visual behavior is closely linked to a number of cognitive processes of visual perception (Rayner 1995). In the first study of its kind, Bulling et al. (2011a, b) demonstrated that they could automatically recognise visual memory recall from eye movements In another study, Tessendorf et al. (2011) showed that high cognitive load during concentrated work could be recognized from visual behavior with high accuracy. Finally, Bednarik et al. (2012) investigated automatic detection of intention from eye movements.

(4) Passive eye monitoring is useful for diagnostic applications in which the user's visual behavior is only recorded and stored for later offline processing and analysis with no immediate reaction or effect on the user's interaction with the world (Duchowski 2002). While passive eye monitoring has traditionally been conducted in laboratory settings a current trend is to move out from the laboratory and to study people in their natural, everyday settings. For example, Bulling et al. (2013) proposed and implemented passive long-term eye movement monitoring as a means for computing systems to better understand the situation of the user. Their system allowed to automatically detect high-level behavioral cues, such as being socially or physically active, being inside or outside a building or doing concentrated work. This information could, for example, be used for automatic annotation and filtering in life logging applications. More information of passive eye monitoring and its applications can be found in Hammoud (2008).

4.1.8 Mobile eye tracking technology

The **mobile eye tracking** device is on the one hand very precise and easy to handle in the calibration procedure which must forego any use of an eye tracking device. However, this device is mostly restricted to use for Microsoft operating systems, such as, the MS Surface tablet PC, and not able to handle Android devices which cover the majority of the market of tablet PCs. Furthermore, the attachment to the tablet might cause problems in the usage for persons with dementia, and they might tend to use it not that often. The specifications of the Tobii EyeX mobile eye tracking device are outlined in detail in Table 1.

(a) (b)

Figure 4. Mobile eye tracker solution Tobii EyeX. (a) The mobile eye tracking device is easily attaches to tablet PCs. (b) The mobile eye tracking device in operation while a person with dementia is performing a playful cognitive control task.

Table 1.Specification of the Tobii EyeX[1] mobile eye tracking device.

Sampling rate	➢ 60 Hz
Accuracy	?
Spatial Resolution	?
Latency	15 ms +/- 5 ms
Calibration	?
Operating range	18-40" / 45-100 cm
Tracking region	16 x 12" / 40 x 30 cm at 26" / 65 cm
Screen size	Up to 27"
API/SDK	EyeX SDK for C/C++
Software	Tobii EyeX Software included
Output data	Gaze coordination stream
Weight	/ 91 grams
Connectivity	USB 3.0
Costs	119€
Operating system	Windows 10, Windows 8.1 or Windows

[1] http://www.tobii.com/xperience/

4.2 VR based eye tracking

While most prior research used eye tracking sensors for interacting with desktop monitors, recent advances in head-mounted displays (HMDs) for Virtual Reality (VR) have also driven development of head-worn eye trackers. VR HMDs with eye tracking technology are becoming more accessible, such as the Tobii HMD (Tobii VR, 2018) or the FOVE HMD (Fove VR, 2018). Using an HMD with such capability, a computer can observe and learn user attention. Well-designed eye gaze-based interaction could potentially offer more natural and implicit interaction that impacts the VR experience in a significant way.

Virtual Reality (VR) is a promising field for new forms of therapy. It is a flexible tool for therapists to provide an immersive and interactive environment for patients. VR has been used in the treatment of phobias (Bouchard et al., 2006), therapy of addictions (Gamito et al., 2014) and motor rehabilitation (Archambault et al., 2014; Tsirlin et al., 2009). Implementing eye-tracking into head-mounted devices (HMD) introduces a precise real-time assessment of attention and opens up new possibilities for VR-based therapy, e.g. analyzing cognitive processes (Holmqvist et al., 2011; Vidal et al.. 2012) or diagnosing medical conditions (Tsirlin et al., 2009). Additionally to common VR-related issues, specific conditions persist when using HMDs in a clinical context. Strict hygiene requirements have to be met to protect patients, particularly considering the prevalence of multi-resistant organisms.

Adding eye-tracking to VR therapy enhances the possibility to accurately measure attention in real-time. Especially in therapy methods where the focus of attention plays a major role, e.g. mirror therapy for stroke patients (Ramachandran et al., 1995) or treatment of social anxiety (Grillon et al., 2006), therapists strongly benefit from an objective measurement of patients' attention. Stimuli can be rendered depending on the focus of attention. Renaud et al. developed an environment which changes when the patients gaze behavior is indicating visual avoidance, as this would decrease therapy effectiveness. The visual field can be masked partially according to the gaze position for rehabilitation of hemispatial neglect (Baheux et al., 2005; Renaud et al., 2011). Furthermore, continuous assessment of attention opens up new potential for realizing adaptive shaping of therapy difficulty, not only based on performance measures but also depending on patients' cognitive abilities at any given moment (Squeri et al., 2011).

Another advantage of eye-tracking in VR is hands- free interaction. Users found gaze based interaction to be easier and more intuitive than pointing in VR (Tanriverdi & Jacob, 2000). This overcomes the challenges of common human-computer interaction (e.g. pressing a key) while wearing a HMD. Interaction methods are predominantly based on eye-fixation time which has been widely adopted for 2D interfaces to solve the Midas touch problem (Jacob, 1990). Fixation or dwell time is a standard delimiter for indicating a user's intention to select an object through eye gaze alone. Dwell time typically ranges from 450 ms to 1 second for novices, but can be improved over time to around 300 ms in the case of gaze typing. However, this time constraint can negatively impact the user experience. For example, when the required dwell time is too short, it puts pressure on the user to look away, avoiding accidental selection, but if it is too long, it results in longer wait times.

4.2.1 Web camera based technology

Despite the availability of accurate, commercial gaze tracker devices working with infrared (IR) technology, visible light gaze tracking constitutes an interesting alternative by allowing scalability and removing hardware requirements. Over the last years, this field has seen examples of research showing performance comparable to the IR alternatives. In the review work of (Ferhat & Vilarino, 2016) a survey on previous work on remote, visible light gaze trackers is given which furthermore analyzes the explored techniques from various perspectives such as calibration strategies, head pose invariance, and gaze estimation techniques. In addition, it also provides information on related aspects of research such as public datasets to test against, open source projects to build upon, and gaze tracking services to directly use in applications.

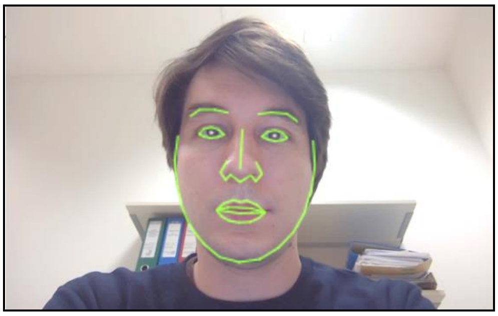
5 Web camera based eye tracking

5.1 Web camera based eye tracking

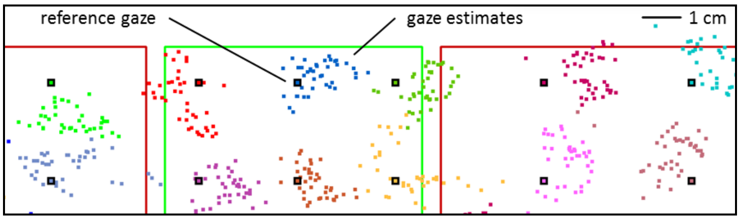
The alternative **web camera based** eye tracking is a state-of-the-art university based solution and it is highly probable that we will see highly improved methods at the market over the next few years and from many sides, as it has been communicated by Apple, Microsoft and SAMSUNG that these companies are working towards a full integration of seamless eye tracking into their desktop devices. However, so far we are using a prototypical solution and are able to benefit from an integrated eye tracking device already during project time.

The implementation of web camera based eye tracking is done in cooperation with J. Huang, Brown Univ., RI, USA, published at Intl. Joint Conference on AI, 2016. The web camera interface can be applied to MS Surface as well as to any Android based tablet PC as well possible. It is based on Java Software and from this is platform implementation independent. Figure 5 depicts features and assessment of the accuracy of the current version of the web camera based eye tracking.

Figure 5a depicts a video frame captured from the MS Surface tablet PC using the web camera based eye tracking device overlaid with information about the automatically determined position of the face in the video frame as well as facial features within the face pose.



(a)



(b)

Figure 5. Web camera based eye tracking. (a) Video frame captured from the MS Surface tablet PC using the web camera based eye tracking device overlaid with information about the automatically determined position of the face in the video frame as well as facial features within the face pose. (b) The evaluation of the accuracy of the web camera has resulted in a precision that is acceptable for the requirements in the frame of project objectives since the methodology of JRD is based on transition statistics between visited areas of interest in the screen layout and not on the precision of the eye tracking device which is estimated to be within a ca. 1.75 cm radius around the ground truth.

Figure 5b depicts the evaluation of the accuracy of the web camera has resulted in a precision that is acceptable for the requirements in the frame of project objectives since the methodology of JRD is based on transition statistics between visited areas of interest in the screen layout and not on the precision of the eye tracking device which is estimated to be within a ca. 1.75 cm radius around the ground truth.

Since the mobile eye tracker from Tobii EyeX is estimated to have a deviation of max 1cm, the accuracy of the web based eye tracking devices is not too far from the mobile eye tracking device.

The system requirements for the web camera based eye tracking device are the following:

- Web camera with min. 1 MPixel video frame size.
- Web camera with min. 20 Hz video rate.

6 Playful gaze interfaces in PLAYTIME

6.1 Attention games with MIRA

MIRA (*Mobile Instrumental Review of Attention*) is part of the PLAYTIME suite of games and particularly contributes in a playful and pervasive manner to the short term indicative **assessment of the psychological functioning** of the user. In the time interval between the clinical screening performed by the neuropsychological experts there is the need to observe the mental state on a short term basis, such as, to alert experts in time in case of rapid mental deterioration. Furthermore, MIRA provides daily and direct **indicative feedback** to the person with dementia and the informal as well as formal caregivers about capabilities in cognitive control that are associated with attentional capacity. The extracted attention features are characteristic for specific aspects in executive functions situated in human prefrontal cortex that are known to be particularly impacted in persons with dementia, such as, the Alzheimer disease. Finally, MIRA potentially serves as a **training playground** by providing a spectrum of perceptual and cognitive stimulation through its interaction patterns that contribute to a wider scope of an intervention strategy.

The eye tracking based serious game solution in MIRA is a direct eye movement sensing, feature extraction, analytics and training component including decision support on several specific aspects of mental state, such as, emotional and cognitive performance. Video camera based eye tracking technology provides the technological basis for the analysis of users' interaction patterns in simple serious games that are designed in the line of well-known neuropsychological tests, such as, the antisaccade test (AST) or Visual Paired Comparison (VPC) task.

Key methodological components of MIRA are displayed in Figure 6. MIRA is an app that can be played on any Android Tablet PCs with a web camera. Its serious game aspect is embedded in simple emotional narratives and implemented via highly usable gaze based interaction patterns. Eye movement data are sensed with the video based eye tracking technology and instrumental attention features that are highly relevant for the characterization of cognitive control performance are extracted. Finally, the statistical analysis is applied by means of classical and machine learning methodologies including prospects of big data analysis on indicative mental health data.

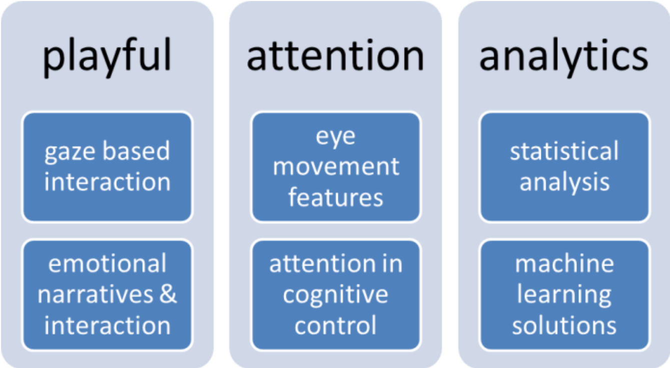


Figure 6. MIRA methodological components for 'Mobile Instrumental Review of Attention'.

In the following, the units for playful tests and training are presented, as well as the queries for emotion analytics.

6.2 MIRA units for playful tests and training

6.2.1 MIA process flow overview

The session initiates with a calibration sub-session to derive the user-specific information for gaze estimates. Then the affective slider page asks the user for its current emotional status. A menu with a list of all available units for attentive game play follows. The attention game is then interactively played using eye gaze for control. Finally, the affective slider page asks the user another time about its current emotional status.

Figure 7. Overview on process flow during a complete MIRA session.

6.2.2 Playful calibration

In a first step, the calibration of the eye gaze with respect to the display plane is adapted with respect to the user position and orientation. A video stream is displayed bottom right about the actual status of the face detection process as well as of the matching quality. Figure 8 shows a screen shot during a specific step in the calibration process where the adaptive face detector has already successfully matched its mask with the face of the user. The border line of the video frame is color coded according to the matching quality: 'green' represents a high quality and red would indicate a low quality. Furthermore, Figure 8 depicts a stimulus (pink flower) with a crosshair which is used by the user to focus the eye gaze on the center of the crosshair and press towards the center with the index finger thereafter.

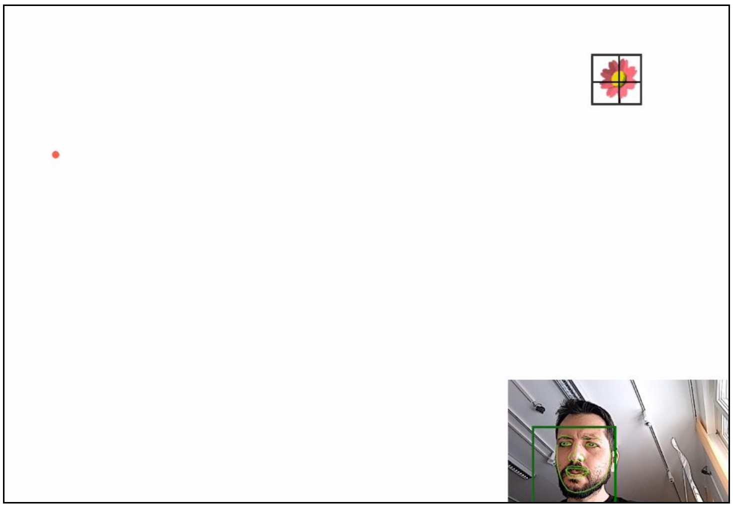


Figure 8. Playful calibration of the eye gaze with respect to the display plane.

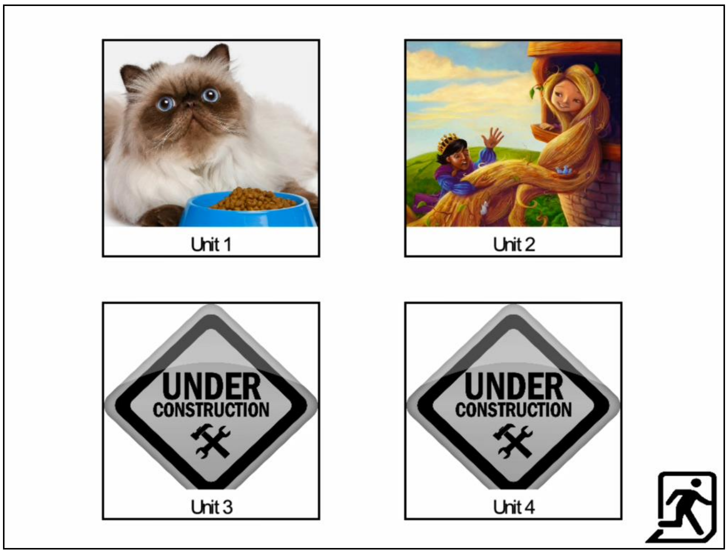


Figure 9. Prototypical menu with options for different attention games, i.e., antisaccade tasks with narratives.

6.2.3 Suite of attention games

The objective of the development of the 'attention component' in PLAYTIME is to provide a suite of attention games for the mental assessment and training of PwD. Figure 9 displays a prototypical menu with options for different attention games, i.e., antisaccade tasks with narratives. The menu can be further extended with other units, such as, for other attention games. In PLAYTIME we focus on (i) antisaccade tests, (ii) visuospatial performanc tests, such as, using 'spot-the-difference' games, and (iii) games that provide data about capacities of actual working memory, such as, 'n-back test' game.

6.2.4 Gamified antisaccade test (GAST)

The gamified antisaccade test (GAST) adds a playful element to the standard antisaccade test in terms of virtual agents that act as mediators between the narrative and the pychological assessment task. These virtual agents take part in the narrative by means of their affective character on the one hand and bridge into the psychological screening task on the other hand. Each game unit is under a specific theme that drives the narrative.

The affective character of the virtual agents is related to the theme in a way that one character has a positive attitude while the other not. The positive attitude of the character makes it perform a positive action. For example, one of the themes is 'feeding a sweet cat' and the positively attached character is a grandmother that carefully feeds the cat in the game and in that way makes her very happy while the negatively attached character is a robber that robs the food out of the cat's feeding bowl so that the cat will suffer.

These virtual agents prinipally appear randomly in time and spatially only at specific, predetermined portals in the scene. If the user spends a sustained ey egaze on any of these charaters it would be 'activated' in terms of performing an action in the manner of its affective character. A short gaze on the character, conversely, would not trigger any action of a character. In this sense, the user will experience that a sustained gaze on the positive character will consequently activate a positive reward, and to prevent from that sustained gaze will prevent the robber to apply a negative reward in the game. Prevention fro sustained gaze might then even be improved by doing an antisaccade upon appearance of the negative character and positioning that saccade on the area of interest in the opposite direction of the screen.

Figure 10 depicts a prototypical screen of the antisaccade task. It includes two 'portals' left and right that indicate the potential emergence of a new stimulus, the gaze pointer (red) and a status bar including the status of the individual game – in this version, the status of the happiness of the virtual cat that should be fed by the 'good stimulus', i.e., the 'grandmother'.

Figure 11 demonstrates a prototypical screen of the antisaccade task, depicting a negative character in the scene that would trigger an action with negative consequences.

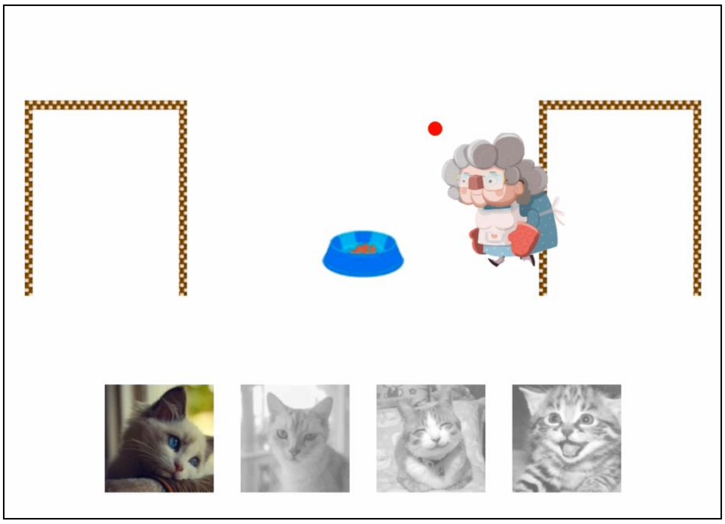


Figure 10. Prototypical screen of the antisaccade task, depicting a positive character in the scene that would initiate an action with positive consequences.

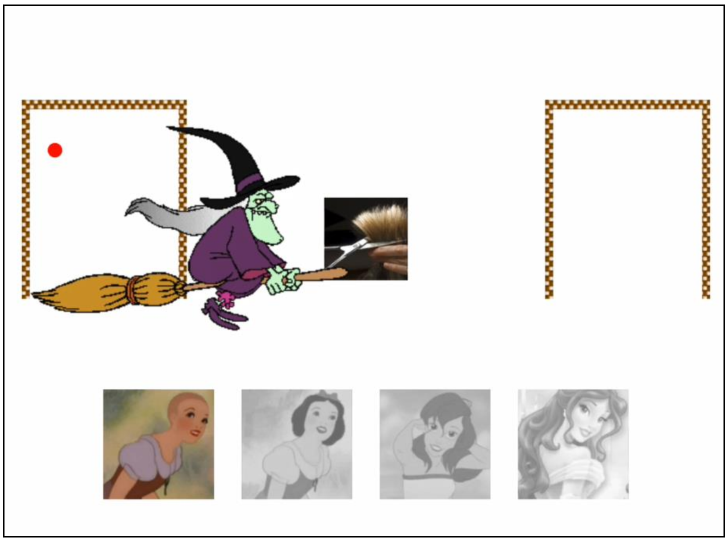


Figure 11. Prototypical screen of the antisaccade task, depicting a negative character in the scene that would trigger an action with negative consequences.

6.3 Gamified gaze based emotion analytics

6.3.1 Affective Slider (AS)

Self-assessment methods are broadly employed in emotion research for the collection of subjective affective ratings. By leveraging on state-of-the-art user interfaces and metacommunicative pictorial representations, (Betella & Verschure, 2016) developed the Affective Slider (AS), a digital self-reporting tool composed of two slider controls for the quick assessment of pleasure and arousal. The AS has two added advantages over comparable implementations: the AS does not require written instructions and it can be easily reproduced in latest-generation digital devices, including smartphones and tablets.

Figure 12 presents the appearance of the "Affective Slider" (AS) that is intentionally displayed using a neutral chromatic palette to avoid bias in ratings due to the emotional connotation of colors. On top of the AS is a written request to input the current emotional value and arousal in German language.

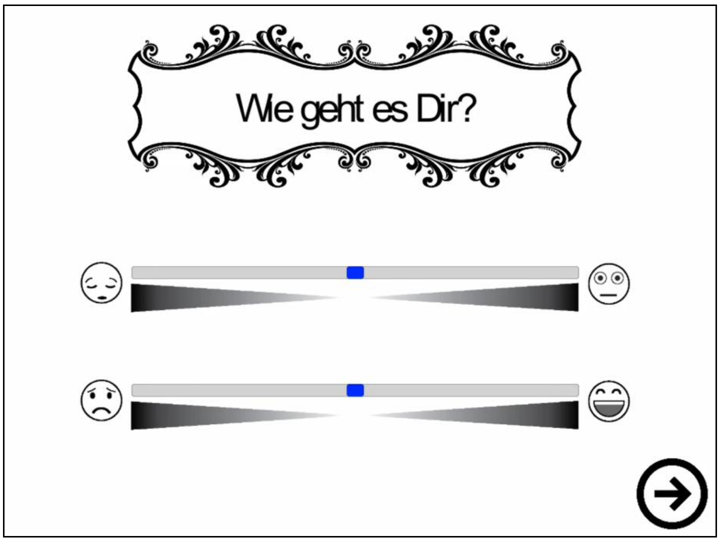


Figure 12. The "Affective Slider" (AS) is a digital self-reporting tool composed of two sliders that measure arousal (top) and pleasure (bottom) on a continuous scale. The AS does not require written instructions and it is intentionally displayed using a neutral chromatic palette to avoid bias in ratings due to the emotional connotation of colors.

6.3.2 Pervasive gaze based emotion screening

Measuring human emotion from affective interaction is an important part of computing for mental health applications. In PLAYTIME, non-obtrusive emotion measurement from eye tracking has been developed in Virtual Reality (VR) technology (Paletta & Dini, 2018) that could be applied as well using the Tablet PC.

In (Paletta & Dini, 2018) PLAYTIME contributed to present a concept and first results of a feasibility study in which affectively weighted imagery of an image database were presented and integrated using discriminative observation, multi-object tracking and video-gaming. Attention preference for affective image classes was measured and relevant correlation with data extracted from a questionnaire on emotional states (MDBF) that would eventually substantiate basic valence classification from eye tracking data was found. The playful approach using the well-known concentration (pairs) game enables frequent repetition of the measurements in mental health care, therapeutic or pedagogical scenarios (Paletta & Dini, 2018).

More details about the approach are described in PLAYTIME deliverable D3.4.1 'Emotion & Psychosocial Analytics'.

7 Gaze analytics

7.1 Eye movement features

7.1.1 Fixation and saccade method

The process of fixation identification—separating and labeling fixations and saccades in eye-tracking protocols—is an essential part of eye-movement data analysis and can have a dramatic impact on higher-level analyses. However, algorithms for performing fixation identification are often described informally and rarely compared in a meaningful way. (Salvucci & Goldberg, 2000) proposed a taxonomy of fixation identification algorithms that classifies algorithms in terms of how they utilize spatial and temporal information in eye-tracking protocols. PLAYTIME uses this taxonomy, that is known to compare well with respect to a number of qualitative characteristics as described in (Salvucci & Goldberg, 2000).

7.1.2 Eye movement features

Based on the fixation detection algorithm described in the previous section, the following eye movement features were extracted:

- Error rate: The percentage of attempts with the error of not doing an antisaccade when one was necessary to do.
- Fixation duration (FD) — The duration of fixations during the trials was collected, and the median fixation duration across the trials was used as an input feature. The use of this feature is motivated by a previously reported significant difference in fixation durations between NC and AD subjects (Scinto et al., 1994). The change in fixation durations is thought to be related to changes in visual spatial attention, saccade initiation, or inefficiency in planning strategy during visual search observed for AD subjects (Ogrocki et al., 2000).
- Re-fixations (RF) — The fixation sequence is used to capture the times when the gaze position re-visits (re-fixates) on previously seen parts of the stimuli. The algorithm detects re-fixation if there are fixations in the proximity of a previously made fixation and the distance between the centers of the two fixations is less than a specified threshold. The “depth” of re-fixation refers to the number of fixations that occurred between the current fixation and the most recent fixation at the same location.
- Saccade orientation (SO)—Saccades were defined by the corresponding endpoints of the fixations. To characterize the saccades, we considered the orientation of the saccades – that is, the angles of individual saccades. For this feature, we considered only the absolute value of the saccade angle, ignoring the direction of the movement (i.e., up or down). The vertical saccades tend to occur within the same stimulus,

whereas others are more likely to move the gaze across stimuli, e.g., switch between different areas of interest in the image.

7.2 Classification of mental state

12 participants with Alzheimer were equipped with a Tablet PC and applied the amicasa training suite used as well the mobile eye tracking units for a period of 6 months (Paletta et al., 2018). The amicasa training suite was configured on a Microsoft Surface Tablet with USB-connected Tobii EyeX mobile tracking technology, providing a 60 Hz sampling of gaze towards the display of the Tablet PC after a 5-point calibration procedure. After a data filtering with a removal of any unit training data with a mean data logging frequency larger than 4.5 Hz, and a removal of sessions where center positions were not fixated after a stimulus presentation, eventually 422 training unit sessions remained for analyses and monthly measurements were aggregated into average measurement values.

Eye movement features were extracted and the prosaccade feature was predictive with respect to CDT (clock drawing task) and MMSE (minimental state exam). The prosaccade (‘PSA’) feature (percentage of gaze within wrong area-of-interest, a kind of weak error rate) was significantly predictive with M=45.7 (S=12.5) for PwD (CDT<5) and M=19.0 (S=12.8) for the PDf (persons without dementia; CDT>=5). Furthermore, the prosaccade feature was as well predictive with M=43.1 (S=12.4) for PwD (MMSE<26) and M=31.8 (S=13.3) for study participants with higher mental status (MMSE>=26).

A support vector machine (SVM) network was tuned with 5-fold cross validation on training data and achieved a classification accuracy of 75% for MMSE (>26) and 100% for CT (>5).

8 Conclusions and Outlook

From the studies described in (Paletta et al., 2018) we conclude that eye movement features can be successfully applied to provide indicators for Alzheimer diagnostics, considering two independent studies that both showed the discriminative power to classify into dementia and non-dementia participants exclusively from gaze data.

Future work will involve larger number of participants in field trials to get more robust estimators for Alzheimer classification. Furthermore, multiple eye movement features will be used for estimation and classification. In addition, multimodal sensing should even lead into better estimates, for example, by incorporating features from movement studies.

9 Bibliography

(**Archambault et al., 2014**) P. Archambault, N. G. Norouzi, D. Kairy, J. M. Solomon, and M. F. Levin (2014). Towards establishing clinical guidelines for a virtual reality arm rehabilitation system, in Replace, repair, restore, relieve - Bridging clinical and engineering solutions in neurorehabilitation, Biosystems and Robotics, 2014, vol. 7, pp. 263–271.

(**Baheux et al., 2005**) K. Baheux, M. Yoshizawa, A. Tanaka, K. Seki, and Y. Handa, Diagnosis and rehabilitation of hemispatial neglect patients with virtual reality technology. Technology and health care: official journal of the European Society for Engineering and Medicine, vol. 13, no. 4, pp. 245–60, 2005.

(**Bandura, 1997**) Bandura A. Self-efficacy: The exercise of control. New York: Freeman; 1997.

(**Bartolini et al., 2005**) Bartolini M, Coccia M, Luzzi S, Provinciali L, Ceravolo MG. Motivational symptoms of depression mask preclinical Alzheimer's disease in elderly subjects. Dement Geriatr Cogn Disord. 2005;19(1):31–6.

(**Berger et al., 2002**) Berger AK, Fratiglioni L, Forsell Y, Winblad B, Backman L. The occurrence of depressive symptoms in the preclinical phase of AD: a population-based study. Neurology. 1999;53(9):1998–2002.

(**Betella & Verschure, 2016**) Betella A, Verschure PFMJ (2016) The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions. PLoS ONE 11(2): e0148037. https://doi.org/10.1371/journal.pone.0148037 .

(**Bouchard et al., 2006**) S. Bouchard, S. Côté, J. St-Jacques, G. Robillard, and P. Renaud, "Effectiveness of virtual reality exposure in the treatment of arachnophobia using 3D games." Technology and health care: official journal of the European Society for Engineering and Medicine, vol. 14, no. 1, pp. 19–27, 2006.

(Boyle & Cahn-Weiner, 2005) Boyle PA, Cahn-Weiner D (2005) Assessment and prediction of functional impairment in vascular dementia. Expert Rev Neurother 4, 109-114.

(**Braver et al., 2014**) Braver TS, Krug MK, Chiew KS, Kool W, Westbrook JA, Clement NJ, et al. Mechanisms of motivation-cognition interaction: challenges and opportunities. Cognitive, Affective, Behavioral Neuroscience. 2014;14(2):443–72

(**Cardinal et al., 2002**) Cardinal RN, Parkinson JA, Hall J, Everitt BJ. Emotion and motivation: the role of amygdala, ventral striatum, and prefrontal cortex. Neurosci Biobehav Rev. 2002;26:321–52.

(Crawford et al., 2005) Crawford, T.J., Higham, S., Renvoize, T., Patel, J., Dale, M., Suriya, A. et al. (2005). Inhibitory control of saccadic eye movements and cognitive impairment in Alzheimer's Disease. *Biological Psychiatry*, 57, 1052–1060.

(D'Onofrioa et al., 2018) D'Onofrioa, G. et al. (2018). Executive Dysfunction Detected with the Frontal Assessment Battery in Alzheimer's Disease Versus Vascular Dementia, ournal of Alzheimer's disease: JAD 62(2):699-711.

(Denckla, 1994) Denckla, M.B. (1994). Management of executive functioning. In: Frames of Reference for the Assessment of Learning Disabilities: New Views on Measurement Issues, Paul H, ed. Lyon GR's, Baltimore, pp. 117-142.

(**Engelman et al., 2009**) Jan B. Engelmann, Eswar Damaraju, Srikanth Padmala and Luiz Pessoa, Combined effects of attention and motivation on visual task performance: transient and sustained motivational effectsFront Hum Neurosci. 2009; 3: 4.

(**Ferhat & Vilarino, 2016**) Ferhat, O. and Vilariño, F. Low Cost Eye Tracking: The Current Panorama, Computational Intelligence and Neuroscience, Volume 2016, Article ID 8680541.

(Fletcher, 1988) W. A. Fletcher, J. A. Sharpe (1988). Smooth pursuit dysfunction in alzheimer's disease, *Neurology* 38.

(**Forstmeier & Maercker, 2008**) Forstmeier S, Maercker A. Motivational reserve: lifetime motivational abilities influence cognitive and emotional health in old age. Psychol Aging. 2008;23:886–99.

(**Forstmeier and Maercker, 2015**) Forstmeier, S., and Maercker, A. (2015). Motivational processes in mild cognitive impairment and Alzheimer's disease: results from the Motivational Reserve in Alzheimer's (MoReA) study, BMC Psychiatry, 2015, 15:293.

(**Forstmeier et al., 2012**) Forstmeier S, Maercker A, Maier W, van den Bussche H, Riedel-Heller S, Kaduszkiewicz H, et al. Motivational reserve: motivation-related occupational abilities and risk of mild cognitive impairment and Alzheimer disease. Psychol Aging. 2012;27:353–63.

(**Fove VR, 2018**) FOVE VR "https://www.getfove.com/ , accessed 2018-02-20".

(**Fratiglioni et al., 2004**) Fratiglioni L, Paillard-Borg S, Winblad B. An active and socially integrated lifestyle in late life might protect against dementia. Lancet Neurol. 2004;3:343–53.

(**Fratiglioni et al., 2007**) Fratiglioni L, Wang HX. Brain reserve hypothesis in dementia. J Alzheimers Dis. 2007;12:11–22.

(**Gabryelewicz et al., 2007**) Gabryelewicz T, Styczynska M, Luczywek E, Barczak A, Pfeffer A, Androsiuk W, et al. The rate of conversion of mild cognitive impairment to dementia: predictive role of depression. Int J Geriatr Psychiatry. 2007;22(6):563–7.

(**Gamito et al., 2014**) P. Gamito, J. Oliveira, A. Baptista, D. Morais, P. Lopes, P. Rosa, N. Santos, and R. Brito (2014). Eliciting nicotine craving with virtual smoking cues. Cyberpsychology, behavior and social networking, vol. 17, no. 8, pp. 556–61, Aug 2014.

(**Gollwitzer et al., 2012**) Gollwitzer PM, Oettingen G. Goal pursuit. In: Ryan RM, editor. Oxford handbook of motivation. New York: Oxford University Press; 2012. p. 208–31.

(**Grillon et al., 2006**) H. Grillon, F. Riquier, B. Herbelin, and D. Thalmann, Use of virtual reality as therapeutic tool for behavioural exposure in the ambit of social anxiety disorder treatment, in Proc. 6th Intl Conf Disability, Virtual Reality & Assoc Tech, 2006, pp. 105–112.

(**Heckhausen & Heckhausen, 2008**) Heckhausen J, Heckhausen H. Motivation and action. New York, NJ: Cambridge University Press; 2008.

(**Holmqvist et al., 2011**) K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, J. Halszka, and J. van de Weijer, Eye Tracking : A Compre- hensive Guide to Methods and Measures. New York: Oxford University Press, 2011.

(**Jacob, 1990**) R.J.K. Jacob. What you look at is what you get: eye movement-based interaction techniques. In: Proc. of CHI 1990, pp. ,1990.

(**Kalivas & Volkow, 2005**) Kalivas PW, Volkow ND. The neural basis of addiction: a pathology of motivation and choice. Am J Psychiatr. 2005;162:1403–13.

(Kaufman, et al., 2010) Kaufman, L.D., Pratt, J., Levine, B., Black, S.E. (2010). Antisaccades: A probe into the dorsolateral prefrontal cortex in Alzheimer's disease. a critical review, *J.of Alzheimer's Disease*. 2010; 19(3):781-93.

(**Kempermann et al., 2002**) Kempermann G, Gast D, Gage FH. Neuroplasticity in old age: sustained fivefold induction of hippocampal neurogenesis by long-term environmental enrichment. Ann Neurol. 2002;52:135–43.

(**Kivipelto & Solomon, 2008**) Kivipelto M, Solomon A. Alzheimer's disease - the ways of prevention. Journal of Nutrition, Health and Aging. 2008;12:89S–94S.

(**Knopman et al., 2003**) Knopman DS, Parisi JE, Salviati A, Floriach-Robert M, Boeve BF, Ivnik RJ, et al. Neuropathology of cognitively normal elderly. J Neuropathol Exp Neurol. 2003;62:1087–95.

(**Kruglanski et al., 2000**) Kruglanski AW, Thompson EP, Higgins ET, Atash MN, Pierro A, Shah JY, et al. To “do the right thing” or to “just do it”: Locomotion and assessment as distinct self-regulatory imperatives. J Pers Soc Psychol. 2000;79:793–815.

(**Kuhl & Fuhrmann, 1998**) Kuhl J, Fuhrmann A. Decomposing self-regulation and self-control: The Volitional Components Inventory. In: Heckhausen J, Dweck CS, editors. Motivation and self-regulation across the life span. Cambridge, UK: Cambridge University Press; 1998. p. 15–49.

(Kuskowski, 1988) M. A. Kuskowski, Eye movements in progressive cerebral neurological disease, in: *Neuropsychology of Eye Movements*, Hillsdale, N.J., c.w. johnston and f.j. pirozzolo, editors edition, 1988, p. 146176.

(Lagun et al., 2011) Lagun, D., Manzanares, C., Zola, S.M., Buffalo, E.A., Agichtein, E. (2011). Detecting cognitive impairments by eye movement analysis using automatic classification algorithms, Journal of Neuroscience Methods, pp. 196-203.

(**Lewin et al., 1944**) Lewin K, Dembo T, Festinger L, Sears PS. Level of aspiration. In: McHunt, editor. Personality and the behaviour disorders. New York: Ronald; 1944. p. 333–78.

(**Maercker & Forstmeier, 2011**) Maercker A, Forstmeier S. Healthy brain aging: the new concept of motivational reserve. Psychiatrist. 2011;35:175–8.

(McPherson & Simpson, 2002) McPherson, E.G., Simpson, J.R. (2002) A comparison of municipal forest benefits and costs in Modesto and Santa Monica, California, USA. Urban For Urban Green, 1, 61-74.

(**Modrego & Ferrandez, 2004**) Modrego PJ, Ferrandez J. Depression in patients with mild cognitive impairment increases the risk of developing dementia of Alzheimer type: a prospective cohort study. Arch Neurol. 2004;61(8):1290–3.

(**Paletta & Dini, 2018**) Paletta, L. and Dini, A. (2018). Emotion Measurement from Gaze in Playful Virtual Reality Interaction, 3rd Symposium on Computing and Mental Health, April 22, 2018, Montreal, Canada.

(**Paletta et al., 2018**) Paletta, L., Lerch, A., Kemp, C., Pittino, L., Steiner, J., Panagl, M. Künstner, M., Pszeida, M., and Fellner, M. (2018). Playful Multimodal Training for Persons with Dementia with Executive Function based Decision Support, Proc. Pervasive Technologies Related to Assistive Environments (PETRA), Corfu, Greece, June 26-29, 2018, ACM Press.

(Perry & Hodges, 1999) Perry, R.J. and Hodges, J.R. (1999). Attention and executive deficits in Alzheimer's disease. A critical review, Brain, 1999 Mar;122 (Pt 3):383-404.

(Pierrot-Deseilligny et al., 2004) Pierrot-Deseilligny, C., Milea, D., Müri, R.M. (2004). Eye movement control by the cerebral cortex. *Curr Opin Neurol*. 2004 Feb;17(1):17-25.

(**Ramachandran et al., 1995**) V. S. Ramachandran, D. Rogers-Ramachandran, and S. Cobb, Touching the Phantom Limb, Nature, vol. 377, no. 24, pp. 489–490, 1995.

(**Renaud et al., 2011**) P. Renaud, C. Joyal, S. Stoleru, M. Goyette, N. Weiskopf, and N. Birbaumer, Real-time functional magnetic imaging- brain-computer interface and virtual reality. promising tools for the treatment of pedophilia, Progress in Brain Research, vol. 192, pp. 263–272, 2011.

(**Robert et al., 2006**) Robert PH, Berr C, Volteau M, Bertogliati C, Benoit M, Sarazin M, et al. Apathy in patients with mild cognitive impairment and the risk of developing dementia of Alzheimer’s disease: a one-year follow-up study. Clin Neurol Neurosurg. 2006;108:733–6.

(**Salvucci & Goldberg, 2000**) Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the Eye Tracking Research and Applications Symposium (pp. 71-78). New York: ACM Press.

(Sjöbeck et al., 2010) Sjöbeck M, Elfgren C, Larsson EM, Brockstedt S, Lätt J, Englund E, Passant U (2010) Alzheimer’s disease (AD) and executive dysfunction. A case-control study on the significance of frontal white matter changes detected by diffusion tensor imaging (DTI). Arch Gerontol Geriatr 50, 260-266.

(**Skinner, 1996**) Skinner EA. A guide to constructs of control. J Pers Soc Psychol. 1996;71:549–70.

(**Squeri et al., 2011**) V. Squeri, A. Basteris, and V. Sanguineti, Adaptive regu-lation of assistance 'as needed' in robot-assisted motor skill learning and neuro-rehabilitation, in Proc. IEEE Interna- tional Conference on Rehabilitation Robotics, 2011.

(**Stern, 2006**) Stern Y. Cognitive reserve and Alzheimer disease. Alzheimer Dis Assoc Disord. 2006;20:112–7.

(**Tanriverdi & Jacob, 2000**) V. Tanriverdi and R. J. K. Jacob, Interacting with eye movements in virtual environments, Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '00, vol. 2, no. 1, pp. 265–272, 2000.

(**Teng et al., 2007**) Teng E, Lu PH, Cummings JL. Neuropsychiatric symptoms Are associated with progression from mild cognitive impairment to Alzheimer's disease. Dement Geriatr Cogn Disord. 2007;24(4):253–9

(**Tobii VR, 2018**) Tobii VR "https://www.tobiipro.com/product-listing/vr-integration/ , accessed 2018-02-20".

(**Tsirlin et al., 2009**) I. Tsirlin, E. Dupierrix, S. Chokron, S. Coquillart, and T. Ohlmann (2009). Uses of virtual reality for diagnosis, rehabilitation and study of unilateral spatial neglect: Review and analysis, Cyberpsychology and Behavior, vol. 12, no. 2, pp. 175–181, 2009.

(**Valenzuela & Sachdev, 2006**) Valenzuela MJ, Sachdev P. Brain reserve and dementia: a systematic review. Psychol Med. 2006;36:441–54.

(**Vidal et al.. 2012**) M. Vidal, J. Turner, A. Bulling, and H. Gellersen, Wearable eye tracking for mental health monitoring, Computer Communications, vol. 35, no. 11, pp. 1306–1311, 2012.

(White et al., 1983) O. B. White, J. A. Saint-Cyr, R. D. Tomlinson, J. A. Sharpe, Ocular motor deficits in Parkinson's disease, *Brain* 106 (1983). 571 –587.