Parameters for Vigilance, Attention and Cognitive Workload within Eye Tracking Recordings

by

Judith Ungewiß
registration number 30090

submitted to

Course of Studies Ophthalmic Optics & Psychophysics
Faculty of Optics & Mechatronics
University of Applied Sciences, Aalen, Germany

in partial fulfillment of the requirements for the degree

Master of Science

First Examiner: Prof. Dr. med. Ulrich Schiefer
Second Examiner: Dr. Enkelejda Kasneci

April 2015

Abstract

Purpose

Usually, a theory of attention upon gazed-at locations is applied. More parameters than gaze location can be derived to improve the theory of attention allocation. The aim of this study was to identify parameters related to eye tracking, that are suitable indicators of attention.

Methods

Binocular eye tracking data was collected with the Eye Tribe tracking system (The Eye Tribe Aps, Copenhagen, Denmark) for the task of visual exploration of the painting *"Unexpected Visitors"* by *Ilya Repin*. 20 subjects (valid data: 19/20) had to look at this painting for about two and a half minutes in order to generate fatigue and inattention. In a second step, suitable parameters of attention were transferred to a data set (8 subjects, valid data: 6/8) on a perimetric task executed with the OCTOPUS 900 perimeter (Haag-Streit, Köniz, Switzerland). Monocular parameters could be applied on the perimetric task, the error rate (false positive and false negative catch trials, 5 % each) were taken as additional parameter.

Results

For the image viewing task, the only parameter showing significance was the average level (a10) of fatigue waves ($p = 0.00024$, ANOVA, $\Delta = -0.8316\ px$). Blink duration ($\Delta = -270.4\ ms$), pupil variability ($\Delta = -0.17868$), saccade length ($\Delta = -0.3135\ px$) and fixation duration ($\Delta = 186.5\ ms$) did not change significantly, but showed relevant trends by differences Δ of their median between the first and last tenth of the recording time. Blink rate and the Index of Cognitive Activity (ICA) did neither show significant changes nor relevant trends. Vergence accuracy failed to indicate fatigue due to variability between subjects and comparatively small effect size. For the perimetric data, in 3 of 6 subjects fatigue waves over a limited time window could be observed. Only for one subject, a relevant increase in false negative responses to catch trials (50 percentage points) could be observed.

Conclusion

Pupil diameter variability, saccade length, fixation duration and fatigue waves were the parameters indicating fatigue. Only the latter parameter has the potential to be applied to perimetric data.

Keywords

eye tracking, vigilance, attention, cognitive workload, visual exploration, perimetry

Zusammenfassung

Ziel

Aufmerksamkeit lässt sich nicht nur anhand von Blickzuwendungen beschreiben. Deutlich mehr Parameter können abgeleitet werden, um die Theorie der Aufmerksamkeitszuwendung zu verbessern. Diese Masterarbeit zielt auf die Identifikation von Parametern, die bei Eye Tracking Experimenten aufgezeichnet, aber teils nur selten ausgewertet werden, ab.

Methode

Bildbetrachtungsdaten für das Gemälde *"Unexpected Visitors"* von *Ilya Repin* wurden mit einem binokularen Eye Tracker (The Eye Tribe Aps, Kopenhagen, Dänemark) bei 20 Probanden (gültige Daten: 19/20) für einen Zeitraum von zweieinhalb Minuten aufgenommen. Geeignete Parameter wurden auf einen Datensatz zur Perimetrie (8 Probanden, gültige Daten: 6/8), aufgenommen mit einem OCTOPUS 900 Perimeter (Haag-Streit, Köniz, Schweiz), transferiert. Die Fehlerrate (falsch positive/falsch negative Fangfragen, je 5 %) wurde eruiert.

Ergebnisse

Aufmerksamkeits-Indikatoren zeigten signifikante Werte für Schläfrigkeitswellen (Average-Level 10, $p = 0.00024$, ANOVA, $\Delta = -0.8316\ px$). Pupillenvariabilität ($\Delta = -0.17868$), Saccadenlänge ($\Delta = -0.3135\ px$), Fixationsdauer ($\Delta = 186.5\ ms$) und Lidschlussdauer ($\Delta = -270.4\ ms$) zeigten keine signifikanten Änderungen, jedoch waren relevante Trends zu sehen, da die Mediane des jeweils ersten und letzten Zehntels der Aufnahmen sich um den Differenzbetrag Δ unterschieden. Die Lidschlussfrequenz und der Index kognitiver Aktivität zeigten keine relevanten Trends. Die Vergenzgenauigkeit eignete sich aufgrund zu hoher Variabilität und zu geringer Effektgröße ebenfalls nicht als Indikator. Während der perimetrischen Untersuchung konnten nur bei 3 von 6 Probanden zeitlich begrenzte Schläfrigkeitswellen festgestellt werden. Nur ein Proband zeigte mit dem Einsetzen von Müdigkeit einen relevanten Anstieg falsch negativer Antworten auf Fangfragen (50 Prozentpunkte).

Fazit

Aufmerksamkeits-Indikatoren sind Pupillenvariabilität, Saccadenlänge, Fixationsdauer und Schläfrigkeitswellen. Lediglich der letztgenannte Parameter kann potenziell auf perimetrische Daten übertragen werden.

Schlüsselwörter

Eye Tracking, Vigilanz, Aufmerksamkeit, visuelle Exploration, Perimetrie

Acknowledgements

This Master Thesis was supported by several people who I would like to acknowledge.

First of all, I want to thank all tested subjects. Without you, I would not have been able to collect any data I needed to provide insights for the purposes of this thesis.

I want to express my deepest gratitude to Prof. Dr. med. Ulrich Schiefer for supervising this Master Thesis. Thank you for being the nicest person I could imagine whenever I asked for help. I always appreciated your constructive criticism, your suggestions for improvement and your advice in all respects.

I also want to thank Dr. Enkelejda Kasneci for co-supervising my study and her efforts involved.

My thanks go to Thomas Kübler for being my first point of contact concerning this work. Your tireless help and support in experimental work, theoretical understanding, paper writing and data presenting was way beyond one could expect and I am really grateful.

Finally, I thank Colleen, Bine and Sonja, who were always prepared to help me to take my mind off my thesis when I needed to, for lots of fun and amusing conversations.

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Aalen, April 2015

Judith Ungewiß

Contents

List of Figures

List of Tables

List of Abbreviations

CFF	Critical Flicker Fusion
cm	centimeter(s)
e.g.	exempli gratia (meaning "for example")
ECP	Evoked Cognitive Potentials
EEG	electroencephalogram
EOG	electro-oculography
Fig.	Figure
fn	false negative
fp	false positive
Hz	Hertz
i.e.	id est (meaning "that is")
ICA	Index of Cognitive Activity
med	Median
min	minute(s)
mm	millimeter(s)
MoCS	Method of Constant Stimuli
ms	millisecond(s)
MSLT	Multiple Sleep Latency Test
MWT	Maintenance of Wakefulness Test
pdv	pupil diameter variability
POG	photo-oculography
PST	Pupillographic Sleepiness Test
PUI	Pupillary Unrest Index
px	pixel(s)
SSS	Stanford Sleepiness Scale
TAP	Test: Alertness Program
TEPR	Task-Evoked Pupillary Responses
VOG	video-oculography

1 Introduction

As eye tracking is becoming more and more affordable and easy to handle, for instance by the release of the The Eye Tribe-eye tracker in 2013, this technology enables several opportunities of eye tracking experiments. Furthermore it is not only possible to collect monocular, but even binocular data.

Usually, a theory of attention upon gazed-at locations is applied. However, many more parameters than mere gaze location can be measured and derived. In addition, gaze location and attention do not always match. A case in point for the difference between gaze location and attention are look-and-find pictures such as "Where's Waldo?" or other ones. These are also called "Wimmelbild" (see Figure 1.1). Looking for a specific person or object requires a lot of attention, although a little number of gazes on the person or object occur.

Figure 1.1: Painting ("Wimmelbild") called *31c3* by *Caro Wedekind*, source: http://foxitalic.de/2014/12/31c3/, license: https://creativecommons.org/licenses/by/3.0/de/legalcode, latest access: 2015-02-13

Within this thesis, the theory of attention allocation with other parameters for attention than mere gaze location is to be improved. A variety of further parameters can be acquired by eye tracking without any additional effort.

1.1 Exposure of the Research Issue

It is well-known that methods for the examination of the vision system based on attention can be boring and exhausting for patients. Two examples for such examinations are exploration behaviour and visual field testing. These are two completely different examinations with different demands, but both are related to vigilance.

A session of static perimetry, for instance, can take up to 15 minutes or even more and all the patients have to do is keeping their gaze stable at a fixation target and pressing a button in case of perception of a stimulus somewhere in the visual field. Therefore, it is - at times - not possible for them to keep up their concentration on the given task.
Results of these examinations are merely usable, if patients concentrate on their tasks and are attentive. For that reason, it is essential to gain knowledge about their attentional status during the examination to be able to decide when an examination session has to be terminated due to the fact, that it does not lead to valid results any longer.

The attentional status of people can also be interesting from other perspectives. The way people observe paintings and direct their attention while doing so, for instance, can gain new art-historical insights.

1.2 Research Objectives

This Master Thesis aims at the identification of parameters, detected by eye tracking, that are sensible indicators for vigilance, attention and cognitive workload. These parameters are compared to each other and validated against pupillography, a method already used for the assessment of patients' attention. Furthermore, it is demonstrated which method for assessing attention seems to be most valuable with regard to specific tasks, i.e. image viewing and perimetry.

1.3 Organization of this Research Project

This thesis starts with a theoretical background of the research issue, including terminology, physiological definitions, and an explanation of applied algorithms. It introduces the state of the art in attention measurement.

Subsequently, the study design ad methodology for an image viewing task is illustrated. Furthermore, a data set of a previously done experiment regarding fatigue during perimetry is presented.

The original data and the derived attention-correlated measures are shown. The performance of parameters in terms of indicating vigilance, attention, and cognitive workload is being assessed. The performance of the indicators is discussed together with their applicability to different examinations. Besides, the limitations of this research project are demonstrated and an outlook with regard to further ideas for the utilization of the gained insights is given.

2 Background

In the first instance, the theoretical background has to be constituted and the state of the art of science and technology has to be described to actually be able to estimate the capability of eye tracking regarding examination methods based on visual attention.

2.1 Vigilance, Attention and Cognitive Workload - Definition and Scope

To make the subject matter of this work clear, it is essential to agree about some definitions, to be clear on the physiology of pupillary osciallation and gaze attention, and to define the algorithms used throughout this study.

Vigilance is derived from the Latin word "vigilantia" and can be translated as wakefulness. As a technical term, vigilance describes primarily the degree of central nervous activation that enables a tested subject to adapt optimally to the current environment (Head 1923).
Used as a medical term, vigilance is used to describe a physiological state of alertness that directly refers to central nervous activation (Canisius and Penzel 2007).
Central nervous activation and the degree of vigilance underlie physiological daytime variation. In the morning and in the afternoon, central activation is generally higher than at night (Wilhelm et al. 2001). For that, indicators of attention used to monitor vigilance also vary during the day (Kraemer et al. 2000).

In contrast, reduced vigilance is called *fatigue* (Weeß et al. 1998). It is defined as a difficulty in initiating or sustaining voluntary activities (Chaudhuri and Behan 2004).
However, the term "fatigue" is not defined consistently in pertinent literature. In general, fatigue is understood as a condition of psychic exhaustion caused by stress or excessive demands. Fatigue is characterized by a subjective feeling of weariness and a reduced achievement potential on physical and cognitive tasks, frequently combined with a lack of interest (Weeß 2006) as cited in (Endres 2009).

A phenomenon often confused with fatigue is *sleepiness* or *drowsiness*. Fatigue and sleepiness are often seen and applied as synonymous (Shapiro et al. 2002). This is on the one hand because both conditions often appear together, on the other hand, "sleepy" and "fatigued" feel quite similar and can at times hardly been distinguished. Although, fatigue and sleepiness arise on completely different scales.
Sleepiness accords to the degree of vigilance where a characteristic reduction of the central nervous activation and inhibition of the Edinger-Westphal cores occurs. Cerebral structures of the posterior hypothalamus are involved. Typically, sleepiness is caused by a decrease of quantity or quality of sleep during the night and can be characterized by falling asleep directly if possible. This reduced level of vigilance can be determined objectively by using the Pupillographic Sleepiness Test (PST) (Weeß et al. 2000) as cited in (Endres 2009).
The state of fatigue cannot be measured objectively, but only described subjectively (Weeß et al. 2000) as cited in (Endres 2009).

There are various definitions of the term *attention*, and even *visual attention*, that have been refined over time (Carrasco 2011).
William James already stated in his *Principles of Psychology* in 1890 (James 1890):

> "Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state which in French is called distraction, and Zerstreutheit in German."

Bleuler stated in 1916 that attention is the allocation of limited ressources of awareness to contents of awareness, for instance to the perception of the environment or of one's own behaviour and action as well as thoughts and emotions (Bleuler 1916 / 1983).
The substance of that is that attention can be directed, whereas fatigue and sleepiness are global terms. This is, for instance, important for image viewing. Attention can change fast while sleepiness is a slow process.

Over the years, there have been several more definitions brought up and the term attention has been more and more differentiated, according to neurobiological and cognitive aspects. For this thesis we will employ a rather generic definition of attention. We note that attention is related to vigilance, but to be understood more broadly, which is why it is used to describe the cognitive state of probands throughout this study.

In addition, the term *cognitive workload* exists. In cognitive psychology, cognitive load refers to the total amount of mental effort being used in the working memory (Sweller 1988).
Cognitive workload is defined by the relationship between human cognitive resource supply and the demand of a task as well as the amount of information-processing resources used per time. So, cognitive workload can be conceptualised as the degree to which an operator's cognitive and perceptual capabilities are taxed while executing tasks (Wickens and Hollands 2000).

Vigilance, attention, and cognitive workload are examplary applied to a perimetric session, as an examplatory examination:
The occurance of fatigue and particularly sleepiness should lead to a termination of a perimetric session as reasonable results can no longer be expected. Visual attention should be kept in the desired direction, which is straight onto a fixation target. Cognitive workload has to be kept on an even level which should neither be too low nor too high.
Another example to apply attention and cognitive workload on is exploration of images: Visual attention and its spatial distribution are of interest. Locations of high cognitive workload have to be detected as they may be of special importance, for instance for art-historical insights.

2.2 Physiological Parameters for Monitoring Vigilance, Attention, and Cognitive Workload

There are several parameters that can be derived form eye tracking measurements and be used as indicators for fatigue. For monitoring vigilance and attention with eye tracking, the consecutively described parameters can be used as important reference points.

2.2.1 Blink Rate and Duration

A blink is a fast, often automatic and involuntary closure and opening of the eyelids, which essentially provides bearing up the precorneal film for a permanent moistening of the cornea in order to protect the eye from drying-out. Additionally, the eyes can get rid of small particles in this way.

Normal eye blink rates range from 10 to 15 times per minute, which corresponds to one blink every four to six seconds. Mean duration of a blink is between 300 and 400 ms (Moses 1981). Several studies executed in the recent past also figured out quite equal values for the eye blinking rate (Ziemssen et al. 2005), (Bentivoglio et al. 1997), (Barbato et al. 2000).
A study on screen handling found, that blink frequency decreases from 9.7 blinks per minute to 4.3 blinks per minute while operating a workstation (Ziemssen et al. 2005).
With an increase in fatigue, the eye blink rate increases as well (De Padova et al. 2009) and so does blink duration (Tanaka 1999).

2.2.2 The Pupil and Pupillary Oscillations

A physiological pupil diameter ranges from 1.5 mm under photopic conditions to 8 mm under scotopic conditions. Under average photopic conditions, pupil diameter is at a range of about 2-6 mm. Pupil diameter decreases by about 0.4 mm per decade. For that reason, the pupil diameter of elderly people ranges between 4-5 mm (Joos et al. 2003).

Pupillary oscillations within constant environmental conditions are due to the fact, that permanent change of the pupil diameter constantly has to correct the retinal illuminance level. So the pupil shows physiological unrest and pupillary oscillations are the result of this feedback loop, specifically under varying photopic conditions. With less attenuation of this feedback loop (as occuring with increasing fatigue), pupillary oscillations are getting larger (Grünberger et al. 1994).
As a synonym for psychophysiological pupillary oscillations, the term "hippus" is used, though sometimes it is applied for the excrescence of pupillary oscillations (Beatty and Lucero-Wagoner 2000). A novel definition says, that a hippus is an inconstant, spontaneous, bilateral, synchronous, rhythmic constriction and dilatation of the pupil with a large amplitude. Particularly in sleepy people, the size change can be observed. Such large changes in pupil diameter are called fatigue waves (Wilhelm et al. 1999).
For that reason, the terms of pupil diameter and pupillary oscillations are often found in the context of vigilance, fatigue and sleepiness. The sleepier a person is and the less he or she tries to suppress sleep, the shorter the time of initial mydriasis is and the higher and more frequent are the resulting pupillary oscillations that can be detected in young, healthy and vigilant persons under *scotopic conditions*. These waves consist of two components:

- Waves of dilatation and constriction that can last from 4 to 40 seconds and can have amplitudes of up to 0.5 mm.
- Superposed fast and inextended oscillations, i.e. constrictions and re-dilatations of a duration of 0.5 to 1 seconds and with amplitudes of 0.1 to 0.3 mm.

These waves show central nervous activation (Lowenstein et al. 1963).

There is also another theory trying to establish a relation between hippus and vigilance: A characteristic hippus with a mean frequency of 2-3 Hz and amplitudes of many fluctuations up to over 1 mm can be observed in sleepy persons. The origin of these oscillations is yet unknown (Korczyn 1987).

It is known that changes in pupil dilation accompany effortful cognitive processing (Kahneman 1973).
A dilatation of the pupil can be observed in subjects with increased cognitive workload, e.g. subjects instructed to solve a task associated with concentration, such as mental arithmetic. Pupil dilatation refers to the complexity of the given arithmetic task. Calculating the product of two large numbers results in bigger pupil dilatation than for small numbers. Pupils return to their previous size within a few seconds of completing the mental work. Therefore, it can be stated that cognitive workload, which is related to attention, affects the pupil and, in turn, that pupil size over time can indicate cognitive workload (Beatty and Lucero-Wagoner 2000).

2.2.3 Version Movements

The foveola (Fovea centralis) is situated in the center of the retina and is the area of highest visual acuity. Here, the density of light-sensitive receptors, particularly of the cones that are responsible for colour vision is at its maximum level. Therefore, highest visual acuity can only be reached in an angle of about 1° around the fixation position. More peripheral objects are recognized with progressively reduced spatial resolution. The reason for this peripheral blur is the convergence of multiple receptors to one ganglion cell (ratio 125:1). Already, from a deviation of 3° from a fixation point, visual acuity is reduced to 50 %. For that, eye movements spotting the foveola immediately onto the point of interest enable clear vision (Joos et al. 2003).

There are several components of version movements of the eyes, which means movements to similar directions for both right and left eye (see Figure 2.1). The most common version movements are saccades and smooth pursuit.

Saccades are very fast movements (up to 1000° per second (Joos et al. 2003)), that turn the eye to its object of interest. They are released either voluntarily by selective behavior plan while exploring the environment or involuntary, for instance by changes within the peripheral field of vision (Joos et al. 2003).
Saccadic velocity decreases with increasing fatigue, so it can be used as an indicator for vigilance measurement (Russo et al. 2003).

During the period of about 30-40 ms before and up to 120 ms after the beginning of a saccade (for short saccades therefore during the following fixation), visual perception is dramatically reduced (Volkmann 1986). So a 10-fold decrease in contrast sensitivity that begins 50 ms before the onset of the actual eye movement and is maximal at motion onset occurs (Diamond et al. 2000). This is called saccadic suppression. In everyday life, one is not aware of this limitation, but looking in a mirror and looking from the right to the left eye and back, one cannot see the eyes moving (Joos et al. 2003).

Under laboratory conditions, after a position change of a target, the central nervous system of an adult person responds with a saccade after a latency of approximately 200 to 250 ms. Several processes are believed to take place during this latency period, such as shift of visual attention to the new target, disengagement of oculomotor fixation, and computation of the metrics of the movement. These processes involve activation of a large circuit of cortical areas, including the parietal cortex and the frontal lobe. Thus, latency of eye movements can be seen as a cognitive-physiological parameter and is expected to increase with decreasing vigilance and attention (Yang et al. 2002).

When looking at a moving object the eyes follow this movement. So, if the object is not too fast and saccades are required, fixation and thus foveation can be kept up by smooth pursuit movements (Joos et al. 2003).
During fixations, as the eye is in relative rest to an object, visual information is gathered. Minimal fixation duration is in general about 100 ms. This value seems feasible because of saccadic suppression. At shorter fixation durations, no information can be perceived (Joos et al. 2003).

Cognitive fixations, fixations used to generate visual perception, last about 150 - 900 ms. Fixations with a duration of more than 900 ms are called overlong fixations ("staring"), fixations with a duration of less than 150 ms are called express-fixations. Both overlong fixations and express-fixations are non-cognitive fixations. With the incidence of fatigue, a higher percentage of overlong and express-fixations is expected (Schleicher et al. 2008).

2.2.4 Vergence Movements

In contrast to version eye movements, during vergence movements the right and left eye are not guided synchroneously into the same direction. Vergence is needed to display objects in varying distances on the foveae of both eyes simultaneously. With gaze changing between objects in different distances, the eyes have to move contrarily (see Figure 2.1). Vergence movement is linked to accommodation in order to achieve a sharp retinal image of objects in any distance. Phylogenetically, vergence movements are comparatively young. They are relatively slow and are disrupted by alcohol and fatigue (Joos et al. 2003).

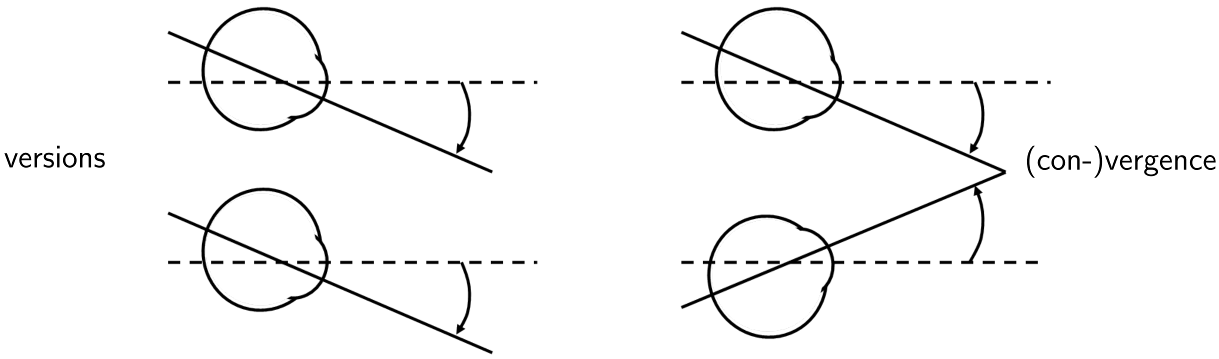

Figure 2.1: Version movements (right and left eye are parallel guided) and (con-)vergence movements (right and left eye are contrarily guided)

Seventy to eighty percent of adult people have heterophoria, what means that the directions that the eyes are pointing at rest position, when not performing binocular fusion, are not consistent with each other (Kaufmann 2004a). Heterophoria appears more pronounced in wearing situations. So, for instance, heterophoria increases with fatigue or appears while dealing with an unfamiliar task (Kaufmann 2004b).

In this study, only horizontal deviations are examined, as vertical deviations are not assumed to provide insight into vigilance and attention monitoring.

2.2.5 Gaze Accuracy

Binocular coordination of eye movement cannot be seen as always perfect. Gaze accuracy names the ability of both eyes to gaze at one target exactly. When gazing at visual targets, for instance luminous points, the abducting eye performs a larger and faster movement than the adducting eye at the start of a saccade. These differences are small but measurable. They cause a transient divergent disconjugacy when beginning and a reverse convergent disconjugacy at the end of the saccade (Vernet and Kapoula 2009).

Vergence errors of up to $\pm 2°$, without diplopia, are common in subjects with normal binocular single vision. Errors of 5° are rare but present (Cornell et al. 2003).
In this study we test the theory that binocular gaze accuracy could decrease with increasing fatigue.

2.3 Algorithms

As there have been several studies to monitor vigilance in the recent past, algorithms to determine basic indicators for vigilance do already exist. Subsequently, two particular algorithms used to determine the pupil diameter and accuracy of binocular vergence will be described, as these algorithms seem very helpful to monitor vigilance specifically in examinations based on visual attention. In addition the wavelet transformation used to extract fatigue waves out of pupillary diameter recordings over time is presented.

2.3.1 Algorithm for Determining the Pupil Diameter

Pupil diameter can, generally spoken, be determined by pupillography. Pupillography describes all methods to record and evaluate pupil activity by means of changes of pupil diameter. Therefore, continuous recordings of the pupil diameter are carried out over a defined period.

To determine the pupil diameter by pupillography during perimetry, an algorithm has been released by Müller in 2013. The pupillary diameter is recorded during the perimetric session. This can be put into practice by the infrared camera integrated in the OCTOPUS 900® perimeter (Haag-Streit, Köniz, Switzerland). The camera delivers a maximum frame rate of about 20 frames per second with a resolution of 320 x 240 pixels in greyscale.

Because of the black coloured pupil and a comparatively bright environment, the lowest values within the image refer to the pupil. Thus, the pupil can be detected by contrast. Here, an intensity threshold of 20 % (according to 80-100 % black) is used to extract the pupillary signal.

Problems to detect the pupil by contrast can appear, if the tracked iris is very dark or if there is a shadow in the lower area of the iris. This kind of shadow can be caused anatomically, by the shape of the eyelid or the background illuminance of the perimeter cupula. In this case, the threshold has to be reduced.
In addition, when cosmetics such as mascara are used, eyelashes can be misinterpreted as pupil because of their dark colour.
Disturbances appear if an examinee is too tired and a shadow appears in the lower area of the eye through squinted eyes and lid closure.

Because of eyelid closure, the pupil can be covered by the eyelid or eyelashes. In this case, the determined diameter is unnaturally small and the pupil is not circular shaped. Thus, every recorded pupil diameter smaller than 1.5 mm can be treated as a blink.
As the diameter of the pupil cannot be determined in a completely closed eye, a reticle, which is implemented in this algorithm, is ignored in that case. Therefore, errors of two pixels (from measured to real pupil diameter) can occur (Müller 2013).

2.3.2 Algorithm for Determining Binocular Vergence

By a binocular eye tracker, signals of both right and left pupil have been recorded by another team (Jainta and Kapoula 2011). A vergence signal, which is the disconjugate eye movement, can easiliy be calculated as Euclidean distance between the position of the left eye and position of the right eye .
Only x-coordinate distance is used. Y-coordinate distances are not related to vergence movements and therefore neglected.

2.3.3 Wavelet Transformation

Wavelet transformation uses wavelike functions in different shapes known as wavelets. Wavelets are used to transform a signal into another representation which presents the signal information in a more useful form. This transformation of the signal is called the *wavelet transform*. Mathematically speaking, the wavelet transformation is a convolution of the wavelet function with the original signal.

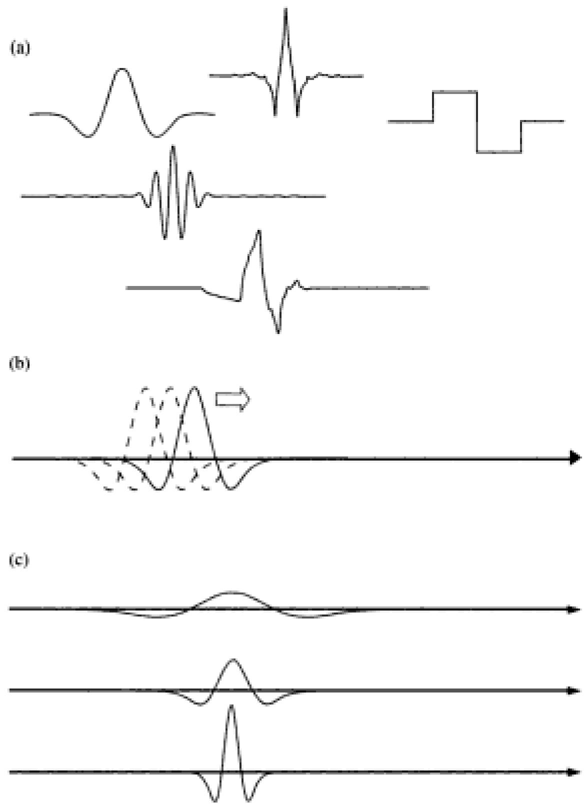

Figure 2.2: Wavelets, (a) some wavelet shapes, (b) location, (c) scale (Addison 2002)

A wavelet can be manipulated in two ways: it can be moved to various locations on the original signal and it can be stretched or squeezed (see Figure 2.2) (Addison 2002).

The wavelet transform contains a pair of filters: One filter is a lowpass filter that implements a scaling function, while the other is a highpass filter. The lowpass filter produces the average signal (a) and the highpass filter produces the detail signal (d) (Weeks 2011).

2.4 State of the Art

Subsequently, the state of the art of eye tracking in general and current efforts to monitor vigilance with visual-based methods, more precisely with pupillography and by determining the accuracy of binocular vergence, are described.

2.4.1 Eye Tracking

The history of eye tracking begins in the late 19th century. The earliest eye trackers were mostly mechanical and difficult to build, not very comfortable for the examinees. In 1898, Huey used a bite-bar with partially cooled sealing wax attached to the mouth piece to ensure examinees would keep their heads still (Huey 1898). Delabarre anaesthetized the eyeball by applying cocaine in order to be able to attach a ring to the eye and to connect that to a mechanical level (Delabarre 1898).Dodge and Cline were the first to use the principles of photography (Dodge and Cline 1901). They took pictures of the reflections of external light sources from the fovea. This has become the dominating technique for recording eye movements in recent years (Holmqvist et al. 2011).

From around 1950, several methods for eye tracking have been developed. The most common methods are, according to Duchowski (2007) the following:

- Electro-oculography (EOG)
  Electro-oculography relies on the measurement of the skin's electric potential differences and was the most common method 40 years ago. Electrodes are placed around the eye. Eye movements are measured relative to head position, so this method is not overall suitable for point of regard measurements unless head position is additionally measured or head movement restricted.

- Scleral contact lens / search coil
  One of the most precise methods to measure eye movement includes the attachment of a mechanical or optical reference object mounted on a scleral contact lens. This method was used, for instance, by Yarbus in 1967 in order to compare scanpaths (Yarbus 1967). The principle method employs a wire coil which is measured moving through an electromagnetic field. Although the scleral search coil is a very precise method (accurate to about 5-10 arc-seconds over a limited range of about 5°), it is not routinely used anymore, as it is also the most intrusive method as the insertion of the scleral lenses requires care and practice.

- Photo-oculography (POG) or video-oculography (VOG)
  These terms embrace a wide variety of eye movement recording techniques including the measurement of distinguishable features of the eyes under rotation and translation. The techniques are different in approach, although they are grouped here. One example from this group is automatic limbus tracking, which involves the use of photodiodes mounted on spectacle frames and the use of invisible, usually infra-red, illumination. Several of these methods require the head to be fixed.

- Video-based combined pupil/corneal reflection
  Although the techniques mentioned above are in general suitable for the measurement of eye movements, they often do not provide point of regard measurement without having the head to be fixed. Video-based trackers use cameras and image processing software to compute the point of regard in real time. The apparatus may be table- or head-mounted and work monocularly or binocularly. These devices are becoming more and more available and are most suitable for use in interactive systems.
  Here, the corneal reflection (known as Purkinje reflection or Purkinje image) of a, typically infra-red, light source is measured relative to the location of the pupil center. Due to the optical construction of the eye, four Purkinje reflections are formed. Typically, the first Purkinje reflection is located by video-based eye trackers. By means of calibration procedures, these eye trackers are able to measure a viewers point of regard. Two points of reference on the eye are needed to seperate eye from head movements. The difference of position between the pupil center and corneal reflection changes only with eye rotation, but remains nearly constant with minor head movements.

Nowadays, eye tracking is conducted with a video-based combined infrared pupil/corneal reflection method, as this is both a comfortable and quite precise method (Duchowski 2007).

2.4.2 Monitoring Vigilance

During the last years, there have been several efforts to monitor vigilance, partially by means of eye tracking, which are mentioned subsequently.

Monitoring Vigilance with Pupillography

The Pupillographic Sleepiness Test (PST) (Wilhelm et al. 1998) was licensed by AMTech in 1997. It is based on infra-red video-pupillography under scotopic conditions. A special software is implemented that includes an algorithm able to detect the pupil diameter, even with distubing factors such as eye movement or eyelid closure. The recorded pupillary oscillation is averaged and declared as Pupillary Unrest Index (PUI) in millimeters per minute [mm/min]. The PST was the first test to prove Lowenstein's assumption, in which the occurance of fatigue waves was related directly to pregressive sleep deprivation. So the PST makes a quantification of fatigue and objective statements about vigilance possible. Today, it is seen as the gold standard in examining sleepiness (Endres 2009).

A quite similar test has been implemented first into a campimetric (Henson et al. 2010), then in a perimetric device (Müller 2013) (see chapter *2.1.3 Algorithms*), both under low photopic conditions. So today it is possible to measure patients' vigilance during examining their visual field by perimetry.

Monitoring Vigilance by Determination of the Accuracy of Binocular Vergence

To date there have not been any efforts to measure vigilance by determining the accuracy of binocular vergence.
However, it is examined, that combined convergent and divergent eye movements are stable under repetition due to saccadic latency. That means, that the *latency* of saccades in vergence eye movement is not related to vigilance (Lang et al. 2014). So far, it is not known, if this is also to be applied to binocular vergence *accuracy*.

Monitoring Vigilance with Combined Methods

There are also some combined methods to monitor vigilance to be mentioned for a comprehensive state of the art, as listed subsequently according to Weeß et al. (2000).

- Multiple Sleep Latency Test (MSLT)
  The MSLT measures latencies of falling asleep and REM phases under polysomnographic conditions. Latency of falling asleep is reduced with increasing fatigue.

- Maintenance of Wakefulness Test (MWT)
  The MWT is a modification of the MSLT. It is an electro-physiological examination method, test criteria are latenciey of falling asleep and REM phases as well.

- Critical Flicker Fusion Test (CFF-Test)
  The CFF-Test is used to determine the visual fusion threshold. It depends on the observation, that intermitting light in a range of frequency lower than 20 Hz is percepted as a flicker signal. By increasing the frequency, the impact of constant light appears continuous upon central nervous activation from a certain, critical frequency on.

- Evoked Cognitive Potentials (ECP)
  Today, this method creates the possibility of illustrating specific reactions of neural structures to a stimulus. So, central nervous activation can be assessed.

- Test: Alertness Program (TAP)
  The TAP is a measurement of reaction time with a warning signal. Via subtest alertness

of the TAP, both tonic and phasic components of the central nervous level of activation can be captured by a computer-based examination.

- Stanford Sleepiness Scale (SSS)
  The Stanford Sleepiness Scale is a method of self-assessment. Patients have to estimate the level of their own vigilance in continuous time intervals on a seven-stage ranking system. The disadvantages of methods of self-assessment like this are, that they only depend in the individual perception of a person. For that reason, methods of self-assessment can hardly be compared to each other, their results are not reliable and, in addition, show accordance to objective methods at rare intervals (Weeß et al. 2000).

In addition, there is the the electroencephalogram (EEG) that is the only physiological signal that has been shown to accurately reflect shifts in alertness, attention and workload (Berka et al. 2007).

2.4.3 Monitoring Attention

In the following, two studies are presented that show how attention can be monitored by eye tracking while the examinees perform different tasks. Attention is monitored by pupillography in one case and by saccadic eye movement metrics in the other one.

Monitoring Task-Evoked Pupillary Responses (TEPRs)

In 1964, Hess and Polt first suggested, that task-evoked pupillary responses might provide a dynamic neurophysiological index of momentary information processing load . They measured pupillary size in five people, who had to mentally calculate the product of two small numbers in four different problems of varying difficulty. Their results were very clear: The pupil of each subject dilatated as each product was mentally calculated. The extend of the observed calculation was almost perfectly monotonically related to the difficulty of the calculation (see Figure 2.3) (Hess and Polt 1964).
Meanwhile, TEPRs are also used to show people's cognitive state concerning perception, memory, responding and language (in terms of translating, shadowing and listening). These studies show, that pupil dilatation can be used for monitoring information processing load and refer to concentration and attention. That is why they can also be used as an indicator for attention (Beatty and Lucero-Wagoner 2000).

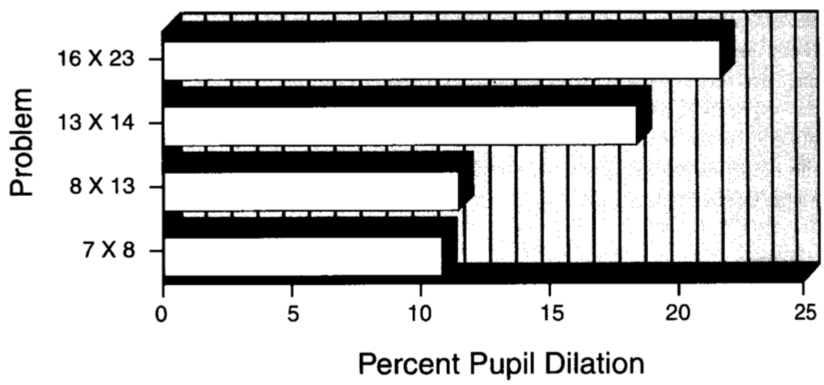

Figure 2.3: Pupillary dilatation during mental multiplications: the original Hess and Polt finding (Hess and Polt 1964)

Monitoring Attention with Saccadic Eye Movement Metrics

In 2013, Di Stasi et al. monitored surgical residents' fatigue levels during their call day using eye movement metrics, objective measures of laparoscopic surgical performance and subjective reports based on standardized questionnaires. They measured the eye movements of twelve members of a resident team (6 males and 6 females) using a head-mounted video eye tracker. It had a similar configuration to a surgical headlight. The performance during three tasks was measured: two simulated laparoscopic surgery tasks (peg transfer and precision cutting) and a guided saccade task. Measurements were performed before and after their call day. Residents rated their perceived fatigue level every three hours throughout their 24-hour shift on a standardized scale based on the Stanford Sleepiness Scale (SSS).

The study found out that time-on-duty saccadic velocity decreased ($p = 0.04$) and subjective fatigue increased ($p = 0.003$), but this did not affect laparoscopic performance. These results support the hypothesis that saccadic indices reflect graded changes in fatigue (Di Stasi et al. 2013).

2.4.4 Monitoring Cognitive Workload

Cognitive workload has an impact on the results of different visual examination methods, especially to methods related to eye tracking. For that, it is explained, how cognitive workload can be measured and its effect on examination results is briefly shown by means of perimetry as an example.

Measuring Cognitive Workload: The Index of Cognitive Activity (ICA)

The Index of Cognitive Activity (ICA) is a method to provide an objective psychophysiological measurement of cognitive workload. The ICA provides an important estimate of the levels of cognitive effort of subjects (Marshall 2000). It is based on changes in pupil dilation that occur as a result of an interaction between a subject and a visual display. The ICA measures abrupt discontinuities in the signal created from continuous recording of pupil diameter. In the presence of effortful cognitive processing, the pupil response occurs rapidly with a reflex reaction of dilatation. At the same time, the pupil shows a reflex reaction to light changes. The ICA separates the light reflex from the dilation reflex.
The index is computed as the number of times per second that an abrupt discontinuity in the pupil signal is detected.
The ICA has been used in a wide variety of applications ranging from simple laboratory tasks to complex user interfaces and was validated against EEG (Marshall 2002).

Cognitive Workload and its Impact on Perimetry

The effect of cognitive workload can also be observed during perimetry: Mental workload reduces the area of one's visual field, and heavy workload leads to a higher reduction of this area than light workload (see Figure 2.4) (Rantanen and Goldberg 1999).

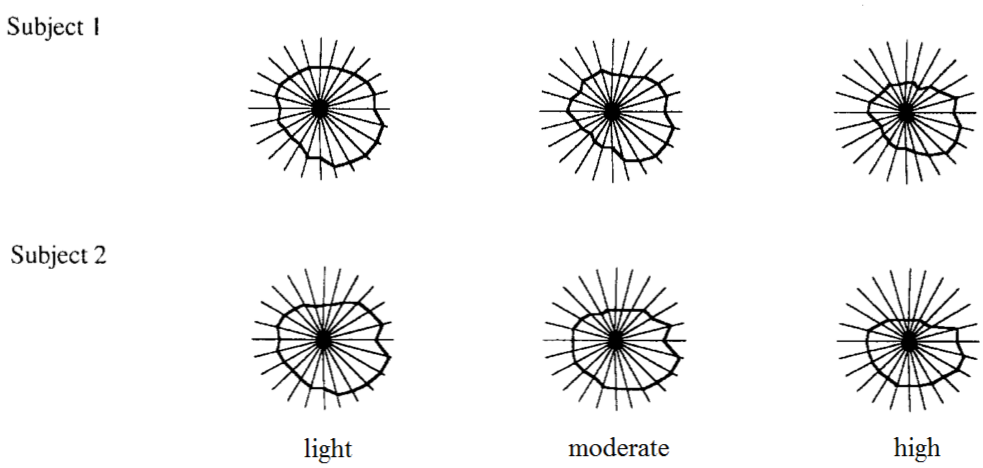

Figure 2.4: The effect of light, moderate and high cognitive workload on the visual fields of two subjects. The vertices extend 90° from the point of fixation. A Goldmann perimeter for kinetic perimetrie with stimulus III 4 e was used. (Rantanen and Goldberg 1999)

3 Study Design and Methodology

In order to identify parameters, that could be sensitive indicators of attention, two data sets are presented: The first one consists of visual exploration of a painting: *Ilya Repin's "Unexpected Visitors"*. This well-established image was the one already used by Yarbus in 1967 for his study on scanpath comparison (Yarbus 1967).

The parameters found to perform well are then applied to the second data set on perimetry that has already been recorded for another study. It is re-evaluated with the parameters that have been figured out to work well as an indicator for vigilance, attention and cognitive workload for the first data set.

3.1 Collecting Data: Visual Examination of Ilya Repin's "Unexpected Visitors"

The data set used to monitor attention of different subjects within this study was recorded together with other data in the context of a different experiment on scanpath analysis. Subjects were pseudonymized and the subject-IDs named were assigned to them during these experiments.

Subjects were instructed to explore images under two different instructions on how to explore the images. Four trials were executed in one block (see Figure 3.1). In the first and the last run subjects were briefed to explore *Ilya Repin's "Unexpected Visitors"* (see Figure 3.2) without any specific task (free-viewing). During the second run, subjects had to explore another painting, during the third run, subjects had to explore *"Unexpected Visitors"* with the task of estimating the age of people in the painting.
Data collected during the fourth run was determined to be recorded exclusively for this study and so is to be evaluated.
Decreasing workload and increasing vigilance with viewing duration are assumed.

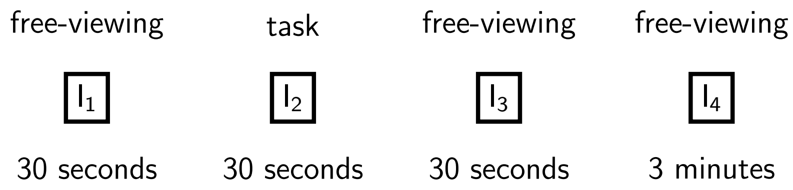

Figure 3.1: Sequence of tasks and durations for the complete eye tracking experiment

Figure 3.2: Ilya Repin: Unexpected Visitors (1888), source: http://www.ilyarepin.org

3.1.1 Subjects

20 subjects (age range 20 - 57 years, 6 male and 14 female) recruited from students and staff of Aalen University of Applied Sciences participated in the study. All subjects gave informed consent to participate in the study. Valid data was collected in 19 out of 20 subjects. For one subject recording quality was insufficient. Subject-IDs were assigned in the context of a larger experiment, therefore subject-IDs larger than 20 were assigned.

3.1.2 Experimental Setup

The experimental setup for this study consisted of a table and chair, a forehead and chin rest, a monitor (FUJITSU Display B24W-7 LED, screen size of 51.5 x 33 cm, resolution 1,920 x 1,200 pixels) to present the image (1,196 x 880 pixels) to explore and an eye tracker (specification see following paragraph). The distance from observer to monitor was 80 cm, the distance from eye tracker to observer was 70 cm (see Figure 3.3).

The Eye Tribe Tracker (The Eye Tribe Aps, Copenhagen, Denmark), an eye tracking system that assesses the location at which a person is looking by means of information extracted from the person's face and eyes, was used. The system is able to estimate the eye gaze coordinates with respect to a screen the person is looking at. The coordinates are then represented as a pair of (x, y) coordinates given in the screen coordinate system. The Eye Tribe Tracker has a sampling rate of 30 Hz and is able to collect binocular gaze data with an accuracy of 0.5 to 1°. This eye tracker in operating with a latency smaller than 20 ms at an operating range from 45 to 75 cm (The Eye Tribe Aps 2013).

The subjects were positioned by a forehead and chin rest. A 9-point-calibration was performed. After that, the subjects were instructed to visually explore the presented image *Unexpected Visitors* by *Ilya Repin*. The presentation time was 155.53 seconds in order to induce a lack of concentration. This value was due to coding and decoding of the timestamps for the frames of the eye tracking recordings. Thus eye movements were recorded over a time interval of 155.53 seconds instead of the originally intended 180 seconds. Subjects were instructed to sit still, not to speak and not to perform chewing motions.

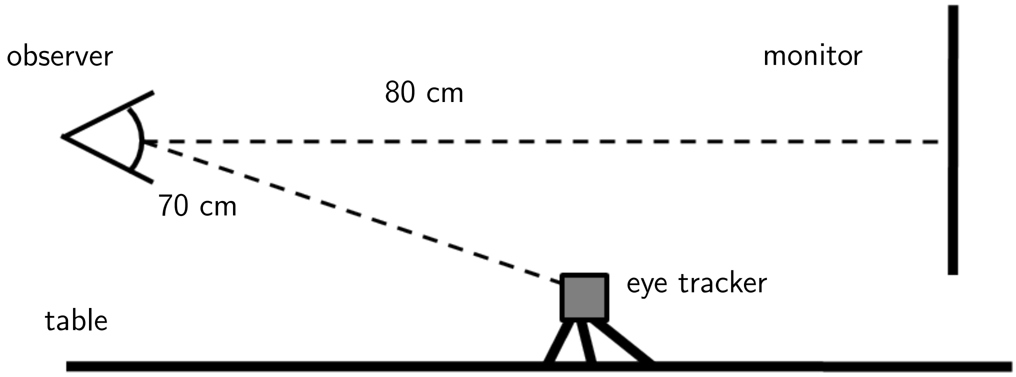

Figure 3.3: Sketch and distances of the experimental setup (profile view)

3.1.3 Evaluation Methods

For the evaluation of the collected data, MATLAB (MATLAB Student Version, Release 2012a (7.14.0.739) 2012-02-09, The MathWorks Inc., Natick, MA, USA) was used.

The following parameters were investigated:

- pupil diameter and pupil diameter variability
- fatigue waves (a10, d10)
- horizontal distance between left and right eye (vergence accuracy)
- blink rate
- blink duration
- saccade length
- fixation duration
- Index of Cognitive Activity (ICA)

For all of these parameters, invalid measurements were filtered and replaced by a linear interpolation from the neighbouring valid data points.
The data is processed as described below. For more detailed information, find the original MATLAB script files on DVD enclosed to this thesis.

Pupil Diameter and Pupil Diameter Variability

The pupil diameter could directly be read from the eye tracking data. A median filter of the order 100 for smoothing was applied.

To calculate the pupil diameter variability, the pupil diameter was extracted from the eye tracking data and scaled down to 1,000 readings. The variance was calculated within the first of overall 100 segments. The variance in the following 99 segments was normalized by the first window. Then, the variability was calculated with a sliding window. After that, a median filter of the order 10 for smoothing was applied.

Fatigue Waves (a10, d10)

Fatigue waves were calculated by wavelet transformation (as presented in chapter *2.1.3 Algorithms*) of the pupil diameter data. MATLAB provided functions for the wavelet transformations in the wavelet toolbox. Average and detail level had to be chosen. The chosen level depended on the frame rate and frequency of the eye tracking recordings. Presently standard values do not exist as there is only insufficient definition of fatigue waves available by now. For different decomposition levels for a and d, see Figure 3.4. For this study, a10 and d10 were chosen as a “best guess”. For these levels most overlap with prior analyses (Henson et al. 2010), (Müller 2013) was found. In addition, they are close to values suggested in pertinent literature (for the rbio3.7 wavelet, see Henson et al. 2010), where a detail level of d8 is suggested for a recording at a different frame rate. Furthermore, detail levels d5 and d7 also showed fatigue waves. For that, these detail levels are displayed and discussed as well.

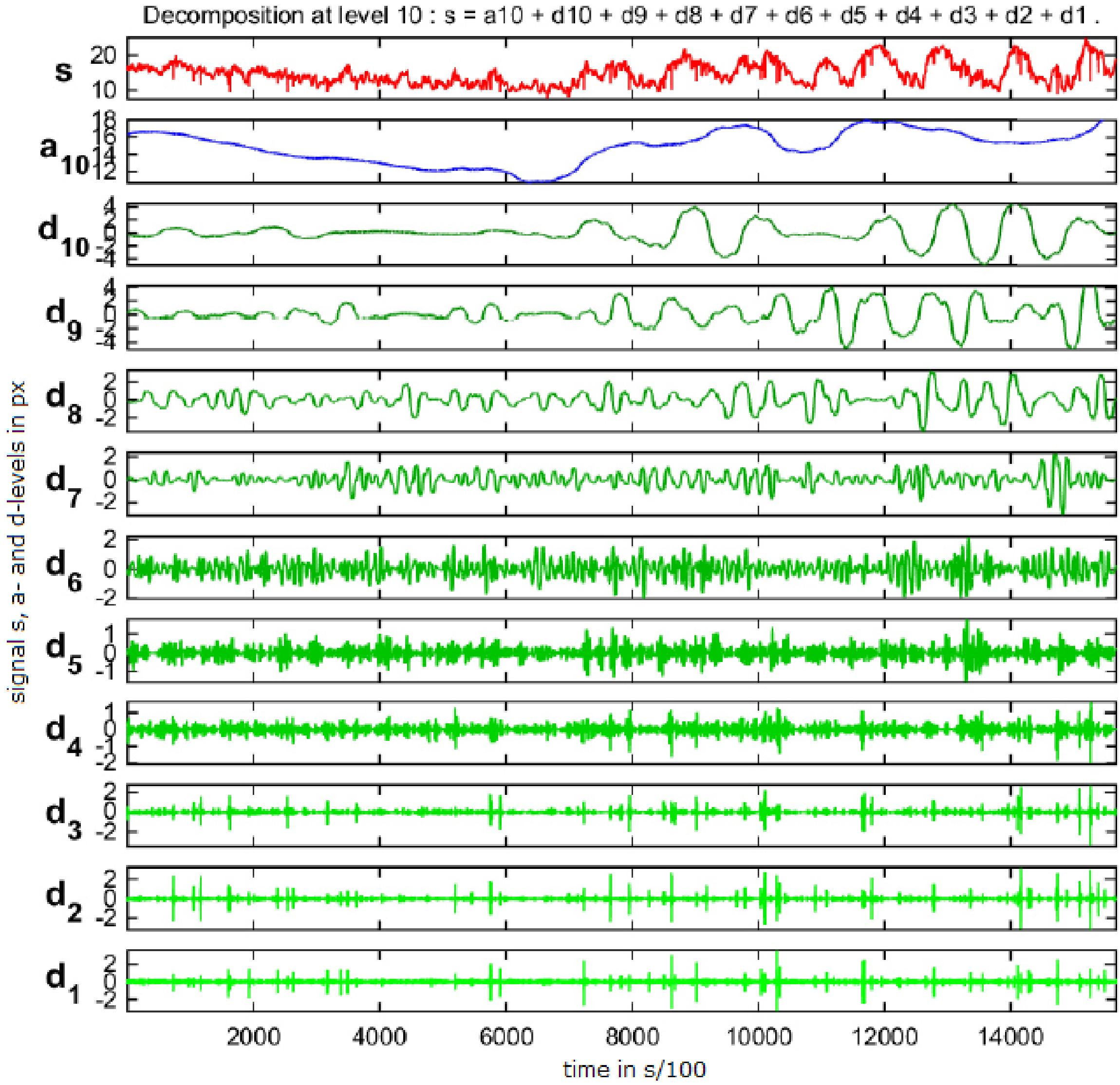

Figure 3.4: Different levels of parameters a and d for fatigue waves. Graphics are calculated with and taken from MATLAB as an example.

Horizontal Distance between Left and Right Eye (Vergence Accuracy)

The horizontal distance between left and right eye, what can be seen as vergence accuracy, was calculated from the eye tracking data. The absolute value was calculated and a median filter of the order 30 for smoothing was applied.

In order to ensure a sufficient determination of the vergence angle, Spearman correlation of vergence accuracy and gaze position in coordinates x and y was tested. If no correlation occured, the determination was rated sufficient and vice versa.

Blink Rate

Blink rate was calculated from the different eye tracking states that were represented as status codes as follows:

4	presence of face detected
7	presence of face + eyes + gaze detected: eyes open
8	tracking failed
16	tracking lost

For that, blinks could be extracted from the eye tracking data by counting eye tracking state 4, expanded with one frame before and after detection. The challenge in blink detection was, that blinks were not actively but passively detected as the loss of eye tracking. This might be due to blink or tracking loss. Eye trackers with face detection as applied partially in this study can be used to distinguish between those. Furthermore the closing movement of the eyelids could be detected as a decrease in pupil diameter. This was done as shown in Figure 3.5.

For the best possible presentation, a moving average over a 15 second duration sliding window was calculated. To compute the number of blinks *per minute*, these moving averages had to be multiplied by four.

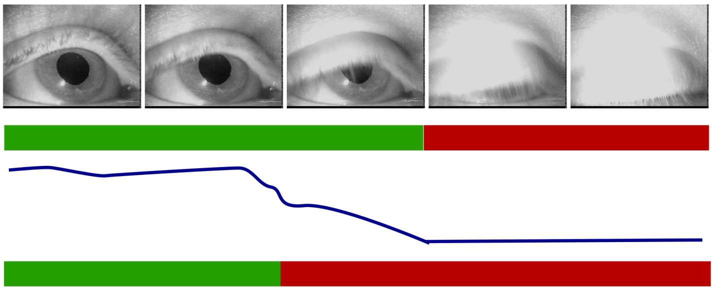

Figure 3.5: Blink detection by evaluation of pupil diameter and validity from eye tracking recording

Blink Duration

As for blink rate, blink duration was calculated from the different eye tracking states (see section *Blink Rate*). The number of frames for one blink was extracted from the data and converted into time units.

Saccade Length

To calculate an average saccade length, fixations and saccades had to be identified. After this had been done, the Euclidean distances between those fixations were calculated.

Fixation Duration

To calculate an average fixation duration, fixations had to be identified. Each beginning and end of the fixations had to be read from their timestamps and were subtracted from each other.

Index of Cognitive Activity (ICA)

The Index of Cognitive Activity was computed as the number of abrupt discontinuities of the pupil signal per second. These detections could be seen by wavelet transformation of the pupil diameter over time (as also seen in section *Fatigue Waves (a10, d10)* and chapter *2.4.4 Monitoring Cognitive Workload*). A median filter of the order 10 was applied.

Parameters were normed in order to compare the *intra*-subject changes instead of the *inter*-subject differences. Figure 3.6 shows the intra-subject differences for subject 12 as an example. First and last tenth of the recording time were compared. During the first tenth of the recording time, different locations were gazed at, whereas during the last tenth of the recording time, staring to only one location occured.

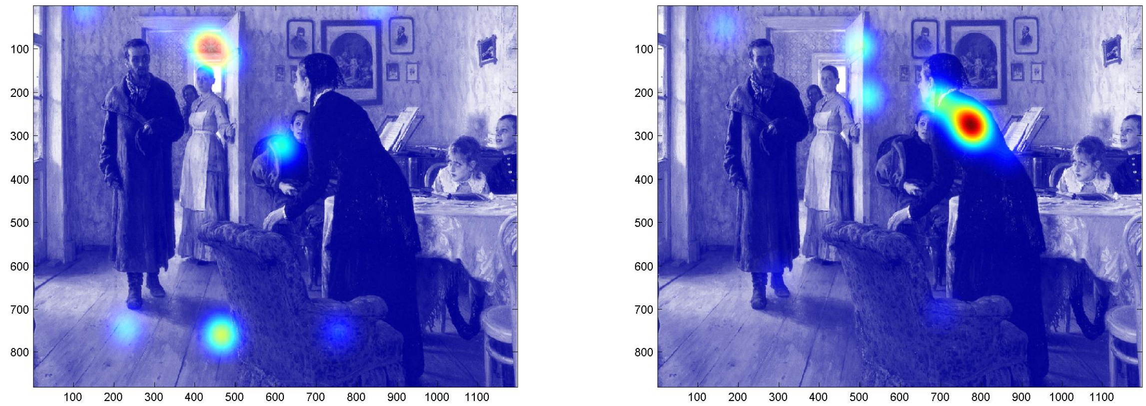

Figure 3.6: Heatmaps for the visual exploration of *Ilya Repin's "Unexpected Visitors"* for subject 12 (left eye) for the first tenth on the left and the last tenth of the recording time on the right side. More frequent and longer fixations appear as warm, shorter fixations appear as cold colours.

For all parameters (except pupil diameter and d10 out of fatigue waves as this is no meaningful visualization for these parameters) the median over all subjects as well as 20^{th} and 80^{th} percentiles were calculated. Therefore, parameters were normed.
Pupil diameter variability (pdv) and vergence accuracy (vac) were normed for every subject separately in order to compare inter-individual differences.

$$pdv_{normed_i} = \frac{pdv_i - med(pdv)}{std(pdv)} \tag{3.1}$$

$$vac_{normed_i} = \frac{vac_i - med(vac)}{std(vac)} \tag{3.2}$$

Another median value as well as another variability for each subject was assumed, as according to the distance from eye tracker to the observing eye, the pupil was represented by a different number of pixels.
a10 was normed to 0 for every subject separately as follows:

$$a10_{normed_i} = a10_i - med(a10) \tag{3.3}$$

For wavelet transformed data, median and peak values for each subject are not that relevant,

whereas variance in that signal is relevant. Hence, a10 was not normed by its variance.
Blink rate, blink duration, saccade length, fixation duration and ICA were not normed in order to show their magnitude in median over all subjects.

To compare two time segments of the recording, which are the first and the last tenth of the recording time, violin plots (with illustration of the median) and an ANOVA for the medians of the two time segments were used. For all parameters, the first and the last tenth of the recording time were compared. This was made because it was assumed that the visual exploration of the painting induced boredom rather quickly. For that, segments had to be separated after the first tenth of the recording, especially for parameters which showed a very small effect or no consistent effect over all subjects. Testing the first tenth of the recording time against the last tenth made sure to actually test the time segment where subject's attention was high and to test two segments of the same extent. Significance level was set to $p = 0.00625$ as eight different parameters were investigated and Bonferroni correction was used. For that, the significance level could be derived from the originally chosen significance level of $p = 0.05$ as follows:

$$p_{crit} = \frac{p_{original}}{8} = \frac{0.05}{8} = 0.00625 \tag{3.4}$$

For comparison of segments and the appropriate statistical testing, parameters were normed. Pupil diameter variability, vergence accuracy and a10 were normed as shown above (for calculation of median and 20^{th} and 80^{th} percentile).
Blink rate (br), blink duration (bd), saccade length (sl) and fixation duration (fd) were normed to 0 as follows in order to be able to compare inter-individual changes, as another median for each subject was assumed.

$$br_{normed_i} = br_i - med(br) \tag{3.5}$$

$$bd_{normed_i} = bd_i - med(bd) \tag{3.6}$$

$$sl_{normed_i} = sl_i - med(sl) \tag{3.7}$$

$$fd_{normed_i} = fd_i - med(fd) \tag{3.8}$$

ICA did not need to be normed as it is an index, which means that an intrinsic normalizations had already been made by the calculation of this parameter.

Not only significant changes but also trends are relevant for attention monitoring. For that, a relevant trend was assumed with the occurrence of a difference between the median values of first and last tenth of the recording time.

All directions were listed with the unit pixels. Translation to a metric unit was resigned, as the parameters were compared inter-individually or shown for combined data of all subjects.

3.2 Re-analyzing Data: Perimetry

Parameters from a previous experiment were transferred to data of a perimetric examination. These data have been collected for another study (Müller et al. 2014) and were re-analyzed.

3.2.1 Subjects

According to Müller et al. 2014, 8 subjects (age range 22 - 60 years, 4 male and 4 female) recruited from students and staff of Aalen University of Applied Sciences participated in this study. Valid data was collected in 7 out of 8 subjects for the *original* study. Due to different evaluation methods, only 6 out of 8 subjects showed sufficient data quality for *this* study. Spheric ametropia ranged from +3.25 to -6.00 dpt, cylindric ametropia ranged from ±0.00 to -1.75 dpt. Else, all subjects were ophthalmogically normal and gave informed consent to participate in the study.

3.2.2 Experimental Setup

Müller et al. 2014 applied MoCS (Method of Constant Stimuli) to assess differential luminance sensitivity within the central visual field with the OCTOPUS 900 perimeter (Haag-Streit AG, Köniz, Switzerland, see Figure 3.7) using the Open Perimetry Interface OPI (Turpin et al. 2012). Stimulus intensity was varied in six steps between 0.50 and 20.08 cd/m^2 with constant background luminance of 10 cd/m^2 and each stimulus intensity level was repeated 30 times in randomized order. Stimulus size III (25.7') was used. Three locations were tested: (0°/0°), (-5°/-5°), (+5°/+5°) (see Figure 3.8). False-positive (no appearance of stimuli but only sound cues) and false-negative (stimuli clearly above previously determined treshold) catch trials were implemented (5 % each). Subjects were instructed to fixate the central target and to press a response button in case of stimulus perception. Pupil diameter was extracted from the built-in infrared camera (frame rate 20 Hz) monitoring the fixation. Stimulus presentation duration was 200 ms and the interstimulus interval started at 1,500 ms and was then adapted to subjects' reation time. So examination duration depended on how fast the response button was pressed in case of perception and ranged from 10 to 14 minutes.

Figure 3.7: OCTOPUS 900 perimeter, source: http://haag-streit-usa.com/octopus-advanced-perimetry/product-specifications/octopus-900.aspx

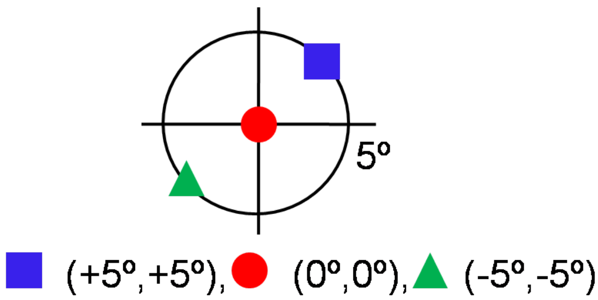

Figure 3.8: Locations of the stimuli

3.2.3 Evaluation Methods

As for the eye tracking data collected during the visual exploration experiment, data was processed as described below. For more detailed information, find the original MATLAB script files enclosed to this thesis on DVD.

For the evaluation of the collected data, MATLAB (MATLAB Student Version, Release 2012a (7.14.0.739) 2012-02-09, The MathWorks Inc., Natick, MA, USA) was used.
The following parameters were investigated and calculated as stated in chapter *3.1.3 Evaluation Methods*:

- fatigue waves (a10, d10)
- pupil diameter variability
- blink rate

Blink rate, pupil diameter variability and a10 were normed to a maximum of 1 (divided by their maxima), and d10 was normed to 0.5 (divided by twice its maximum) for comparability reasons and in order to create well-arranged plots.

In contrast to the previous experiment error rate was taken as an objective parameter for assessing vigilance. The error rate was calculated as follows:
During a perimetric session, 540 questions were asked. Among these 5 % false negative and 5 % false positive catch trials were interspersed. The error rate was calculated as the percentage of wrong responses related to the total number of catch trials. For this, the number of catch trials was enlarged to an array of 1,000 samples, then a moving average over 50 answers was calculated. The error rate was also normed to a maximum of 1 for the individual, respectively. So the value of 0 was excellent with no wrong responses to catch trials, whereas a value of 1 would be the worst case. In addition, every catch trial was plotted as a single spot, including a discrimination between right and wrong responses to the catch trials.
Values up to 20 % for false responses are usually rated as acceptable, values larger than 30 % are rated as inacceptable. Throughout this study, values between 20 and 30 % were also rated as inacceptable.

To investigate the correlation of error rate and parameters a10, pupil diameter variability and blink rate, the Spearman correlation coefficient was calculated. The calculation of p-values was set aside, as these values could only show significance on an individual basis. This was not relevant within this study as there was no prediction for the combined data of all subjects for the reason of inter-individual variations.

4 Results

As shown in chapter *3 Study Design and Methodology*, first data was collected during an examination of visual exploration of Ilya Repin's "Unexpected Visitors". This data set was analyzed to find out which physiological parameters perform well as indicators for vigilance, attention and cognitive workload. As this was done, the gained insights were applied on the previously mentioned data set on perimetry in order to investigate, if the well-performing parameters, figured out for the visual exploration task, are also applicable to the perimetric examinations.

Find the original data on DVD enclosed to this thesis.

4.1 Gaining Insights: Visual Examination of Ilya Repin's "Unexpected Visitors"

In order to identify the most versatile parameters for vigilance, attention and cognitive workload, the parameters mentioned in chapter *3.1.3 Evaluation Methods* were investigated. Subsequently, the results concerning the particular parameters are shown.

4.1.1 Pupil Diameter and Pupil Diameter Variability

Figure 4.1 shows the pupil diameter of the left and right eye of all subjects.

For a better understanding of the changes in the pupil diameter, the pupil diameter *variability* was also examined.
Figure 4.2 shows the median over all subjects as well as 20^{th} and 80^{th} percentiles for the pupil diameter variability of all subjects. This seemed with exception of several peaks, for example after about 65 and 130 seconds, quite even over time. However, the median value of pupil diameter variability was slightly decreasing over time. Since variance between as well as within subjects was large, this effect was masked when only plotting median and 20^{th} and 80^{th} percentiles. A *linear* trend line was applied. It is not known, if a *linear* relation actually

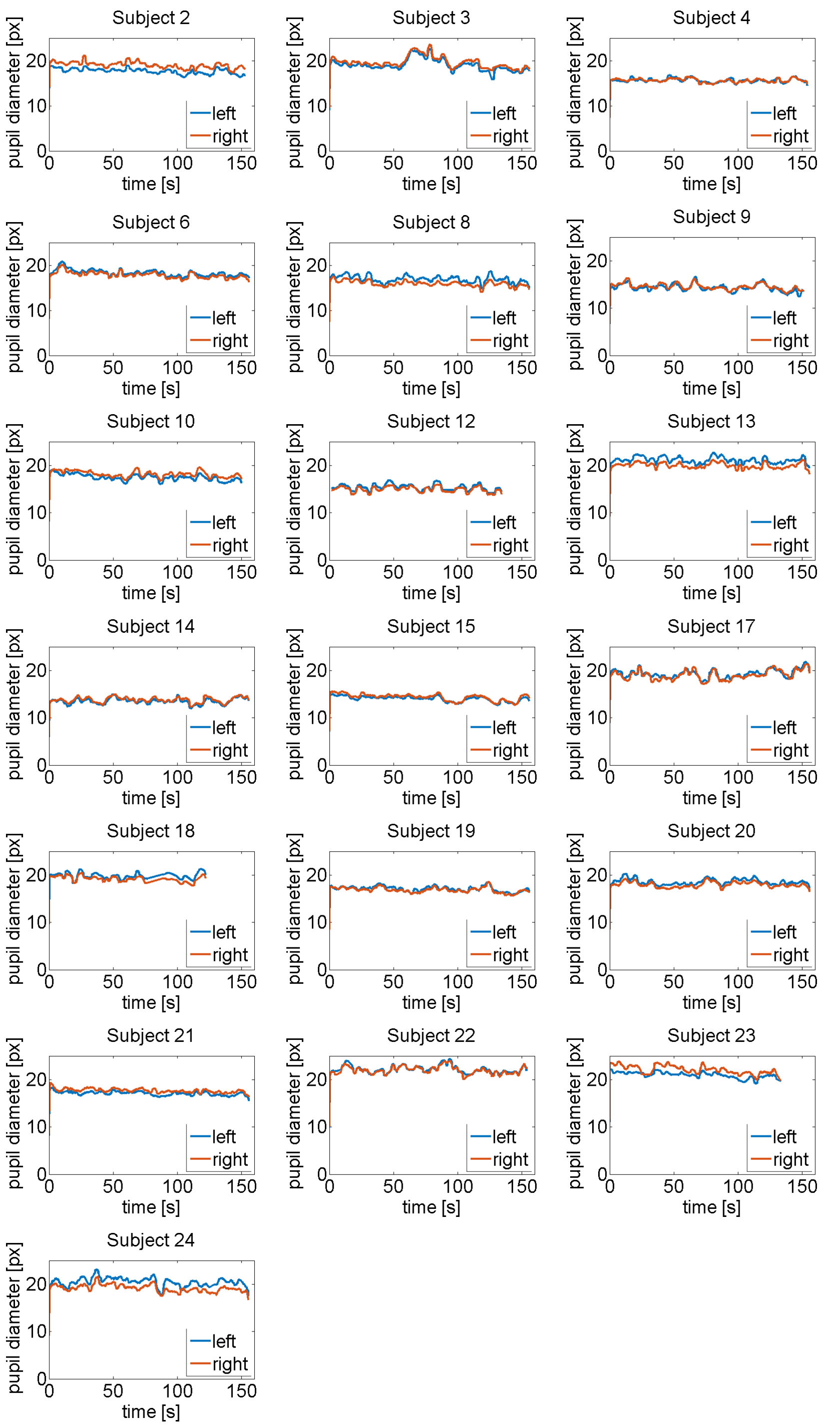

Figure 4.1: Pupil diameters of right (red) and left (blue) eyes in pixels for each participating subject over time in seconds. Subjects 1, 5, 11 and 16 did not participate in this part of the study and subject 7 was excluded because of invalid data.

appeared. Thus, a trend was shown but no data modelling was performed. This trend line with the appropriate equation

$$pdv = -0.00099\ t + 0.11 \tag{4.1}$$

$pdv = normed\ pupil\ diameter\ variability\ [nondimensional]$
$t = time\ [s]$

at 30 frames per second as step size for the time was calculated. Comparing first and last tenth of the recording time, Figure 4.3 shows, that there was no significant difference between these two segments of the recording ($p = 0.0092$, ANOVA). Yet a difference of $\Delta = -0.17868$ for the median of first and last tenth of the recording time appeared (median of the first tenth: $med_1 = 0.05768$, median of the last tenth: $med_2 = -0.121$) and confirmed the decrease over time the applied trend line indicated. For that, a relevant trend was shown.

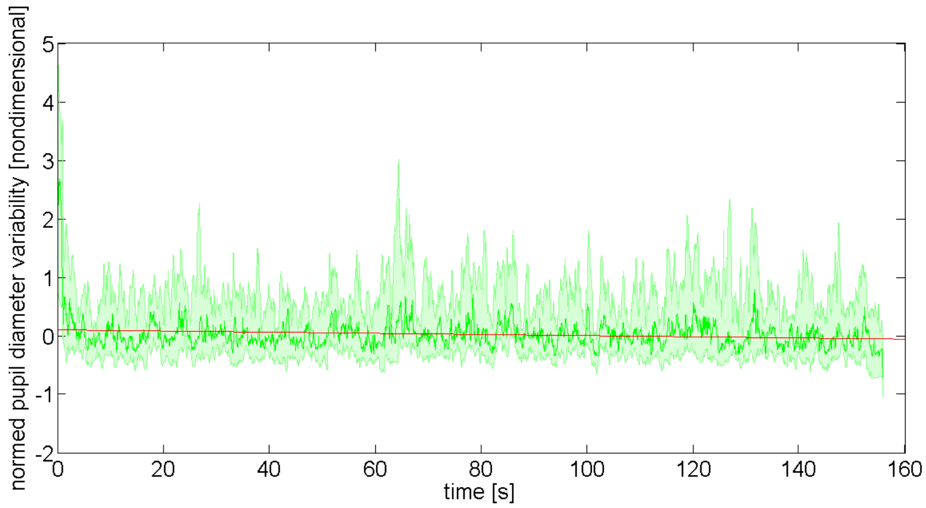

Figure 4.2: Median (50th percentile, dark green) over all subjects as well as 20th and 80th percentiles (light green) for the normed pupil diameter variability (nondimensional) over time in seconds with trend line (red). Pupil diameter variability was normed by subtraction of its median and division by its standard deviation afterwards (see chapter *3.1.3 Evaluation Methods*).

There was a considerable increase in variability towards the end of the recording sessions by trend, which showed, that subjects got unconcentrated and were not as vigilant as in the beginning of the recording sessions. Because of this trend pupil diameter variability seemed to be a usable indicator for vigilance. But under deeper consideration, it could be noticed, that this effect only came up examining the median of pupil diameter variablity over all sub-

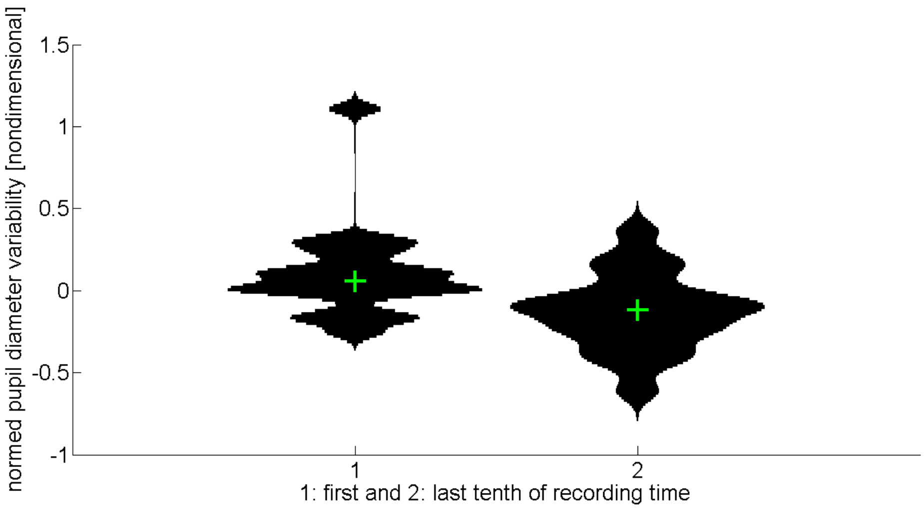

Figure 4.3: Violin plots for comparison of first [0 s;15 s] and last [140 s;155 s] tenth of recording time concerning normed pupil diameter variability (nondimensional, medians are represented by green crosses). Violin plots are similar to box plots but with a rotated kernel density plot on each side. For that, they also show the probability density of the data at different values. Pupil diameter variability was normed by subtraction of its median and division by its standard deviation afterwards (see chapter *3.1.3 Evaluation Methods*).

jects. So for the individual, pupil diameter variability could not be taken as a well-performing parameter for vigilance or attention. Thus only for experiments that examine an average over several subjects, this parameter could work.

4.1.2 Fatigue Waves (a10, d10)

As outlined in chapter *3.1.3 Evaluation Methods*, wavelet transformed data consists of two parameters: a, the parameter for the average, and d, the parameter for the detail level showing the fatigue waves. Figure 4.4 shows the average component for each subject. At the beginning of each recording, the a10 parameter showed a large pupil and, with exception of subjects 3, 14 and 17, even the maximum. This maximum at the beginning of the recording indicated a cognitive workload peak that occured with the task given to the subjects.
Most of the subjects showed overall decreasing average values with evidently smaller fluctuations, although there were exceptions (e.g. subject 17 with an increasing and some subjects like subject 4 and subject 22 with oscillating average waves). In addition, many subjects (e.g. subjects 3, 8 and 20) showed elevations after about 50 seconds.
The average over all subjects is visualized in Figure 4.5. It shows the median over all subjects

as well as the 20^{th} and 80^{th} percentiles. A decreasing median obviously appeared. Besides, this graphic shows an elevation after about 50 seconds. Furthermore, the distance between 20^{th} and 80^{th} percentile (i.e. variability) was larger at the beginning (up to 30 seconds) and at the end (after 100 seconds) of the recording sessions. To illustrate the decreasing a10-values over time, an additional trend line with the appropriate equation

$$a10 = -0.0058\ t + 0.43 \tag{4.2}$$

$a10 = a10 - values\ [px]$
$t = time\ [s]$

was added. As shown, a *linear* trend line was applied. It is not known, if a *linear* relation actually appeared. Thus, a trend was shown but no data modelling was performed.
The decreasing median value of a10 also showed, that there was a workload peak in the beginning of the recording sessions and that workload decreased over time, except the elevation at about 50 seconds. The enlarging confidence interval at the end of the recording sessions showed inter-subject differences in the pupil diameter. A reason for that could be, that different subjects dealt with a boring task in different ways: Some tried to stay concentrated and to discover more and more detail in the painting, others got unconcentrated and began to stare on the picture without further interest.
The elevation most of the subjects showed after about 50 seconds is not explainable by implication. Probably, at that point, subjects realized they were getting unconcentrated and tried to re-concentrate on their task. They started to define "their own task", detected new details of the image and for that, another smaller workload peak occured. Why this should happen at roughly the same time for all subjects is not yet clear and other explanations are possible. There is so far no explanation for the increasing average wave and for that, increasing pupil diameter over time in some individuals (e.g. subject 17).
Comparing first and last tenth of the recording time, Figure 4.6 shows by violin plots that a10 was lower during the last tenth of recording time ($p = 0.00024$, ANOVA). In addition, there was a difference of $\Delta = -0.8316\ px$ for the median of first and last tenth of the recording time (median of the first tenth: $med_1 = 0.6058\ px$, median of the last tenth: $med_2 = -0.2258\ px$). For that, a relevant trend as well as significant changes were shown.

Figure 4.7 shows the detail levels (d10) indicating fatigue waves for each subject. They occured in different shape and with different amplitudes. For instance, subjects 18 and 21

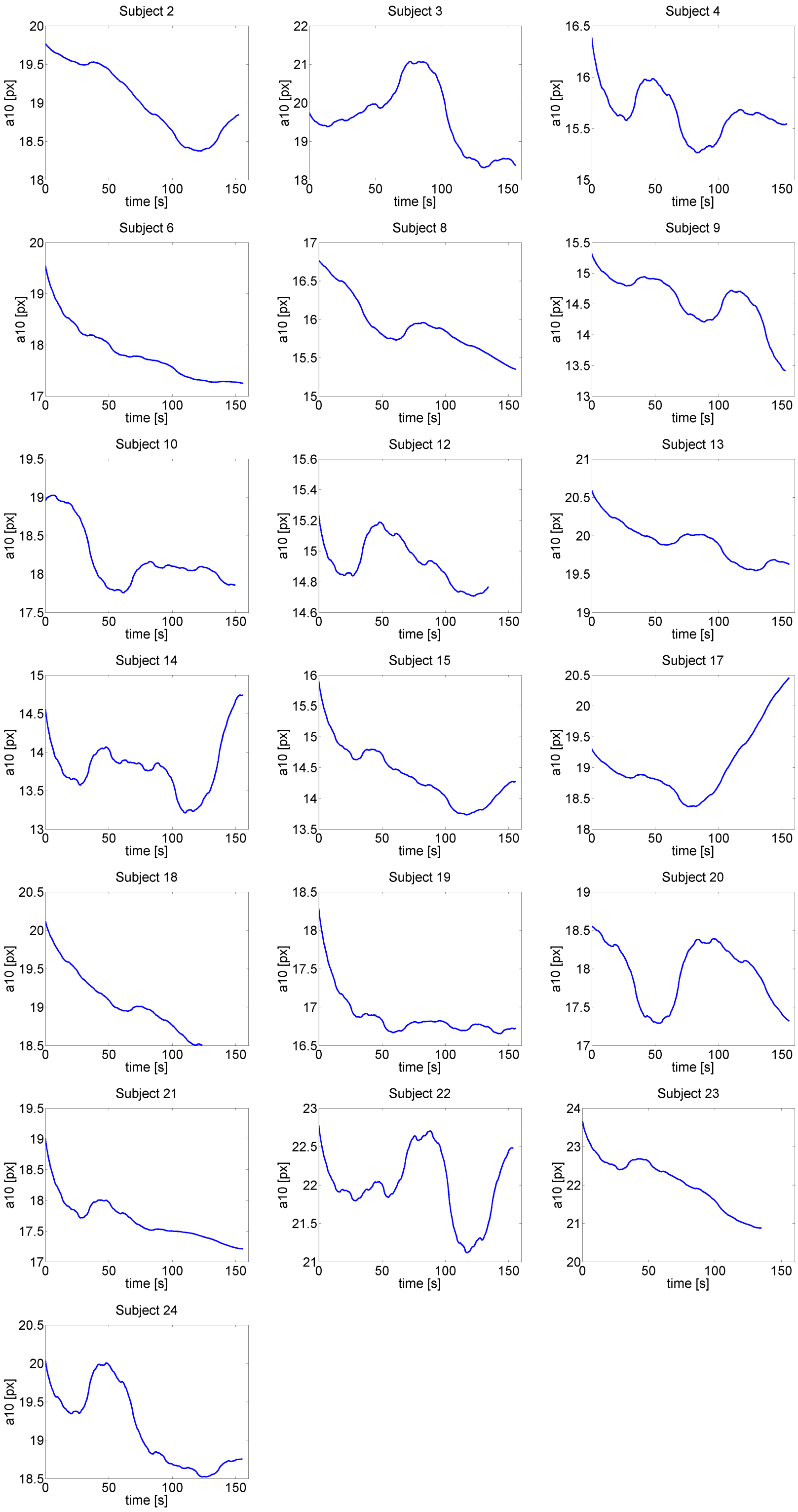

Figure 4.4: Fatigue waves (a10) in pixels for each participating subject over time in seconds. Subjects 1, 5, 11 and 16 did not participate in this part of the study and subject 7 was excluded because of invalid data.

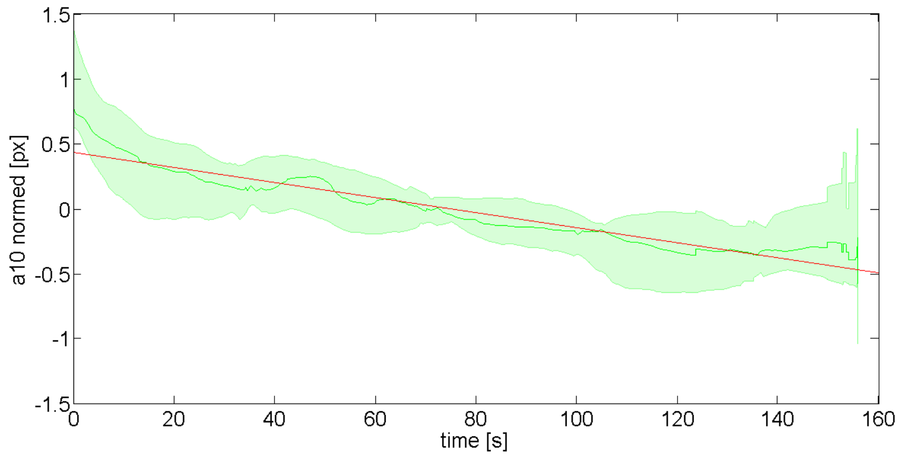

Figure 4.5: Median (50th percentile, dark green) over all subjects as well as 20th and 80th percentiles (light green) for a10 of fatigue waves in pixels over time in seconds with trend line (red). a10 was normed by subtraction of its median (see chapter *3.1.3 Evaluation Methods*).

showed no fatigue waves, whereas subjects 3, 10, 15, 19 and 22 showed fatigue waves. Subjects 4 and 6 showed only temporary fatigue waves.

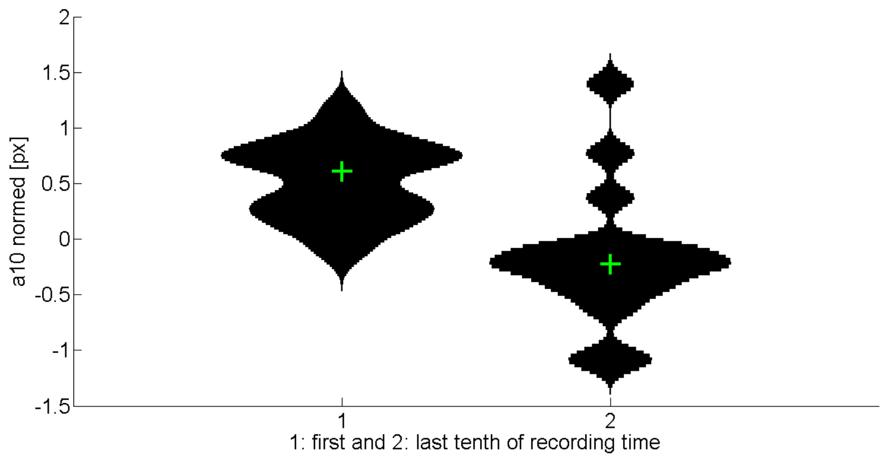

Figure 4.6: Violin plots (as explained in Figure 4.3) for comparison of first [0 s;15 s] and last [140 s;155 s] tenth of recording time concerning a10 (normed) of fatigue waves in pixels. Medians are represented by green crosses. a10 was normed by subtraction of its median (see chapter *3.1.3 Evaluation Methods*).

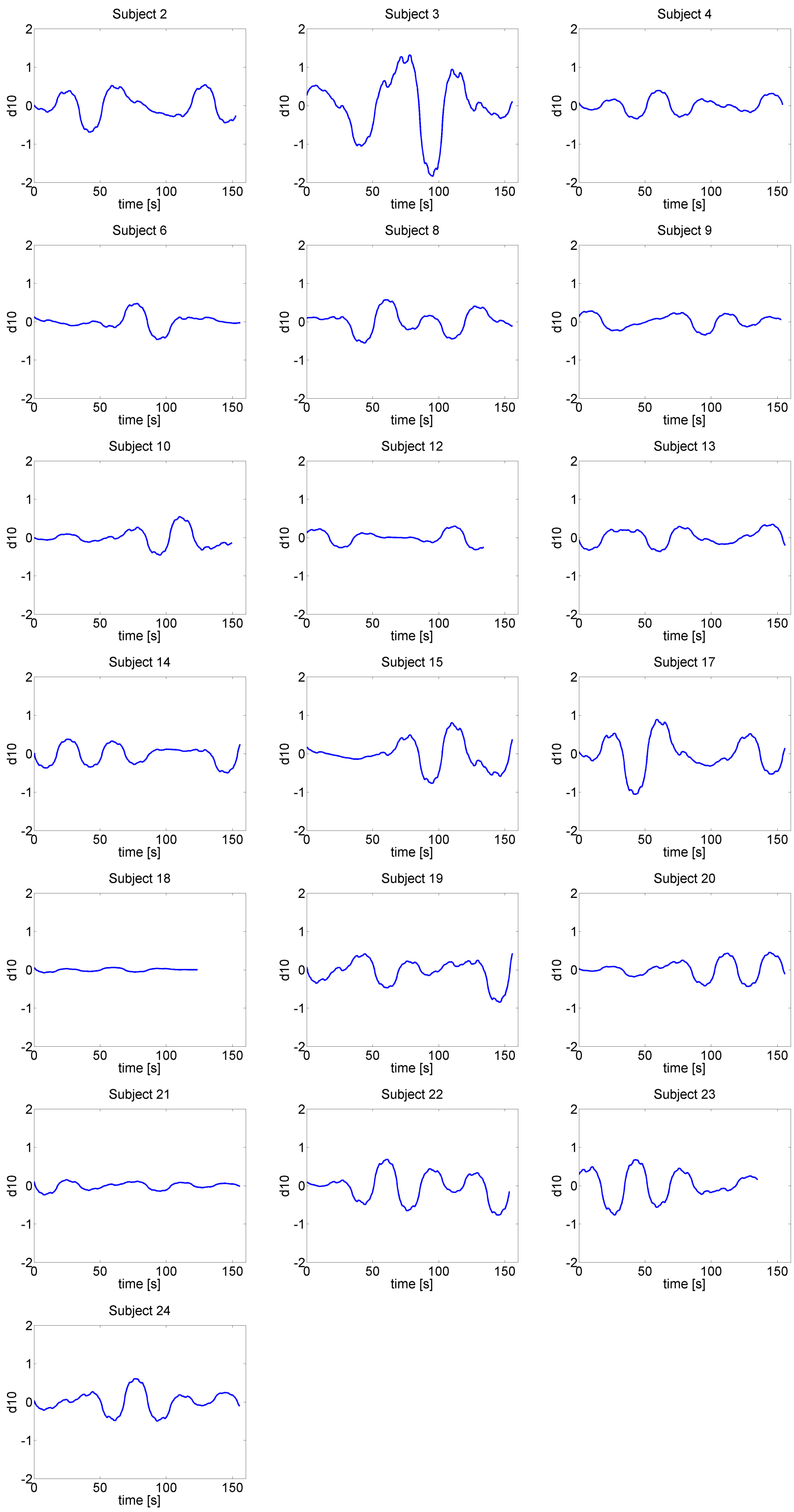

Figure 4.7: Fatigue waves (d10) for each participating subject over time in seconds. Subjects 1, 5, 11 and 16 did not participate in this part of the study and subject 7 was excluded because of invalid data.

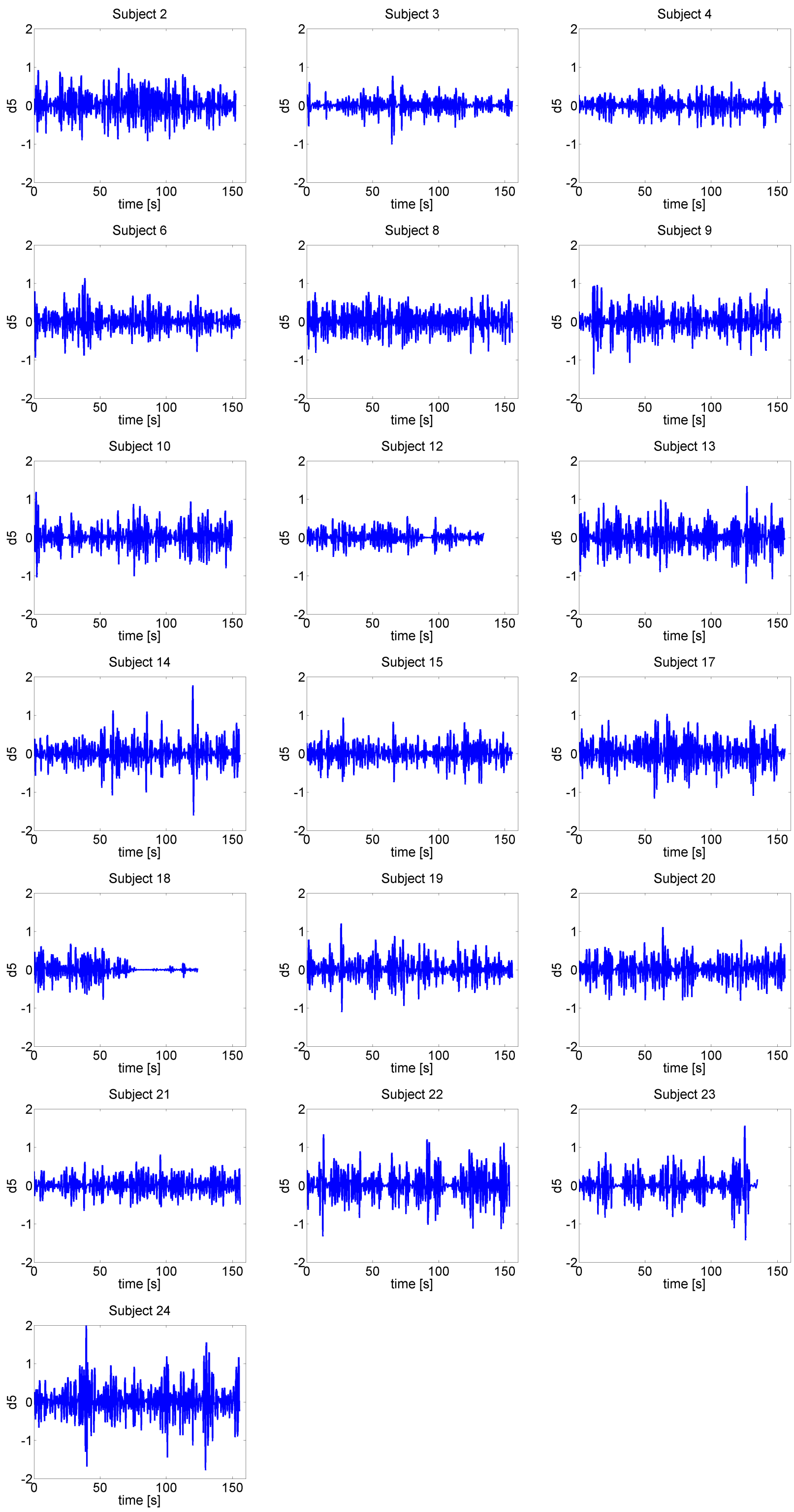

Figure 4.8: Fatigue waves (d5) for each participating subject over time in seconds. Subjects 1, 5, 11 and 16 did not participate in this part of the study and subject 7 was excluded because of invalid data.

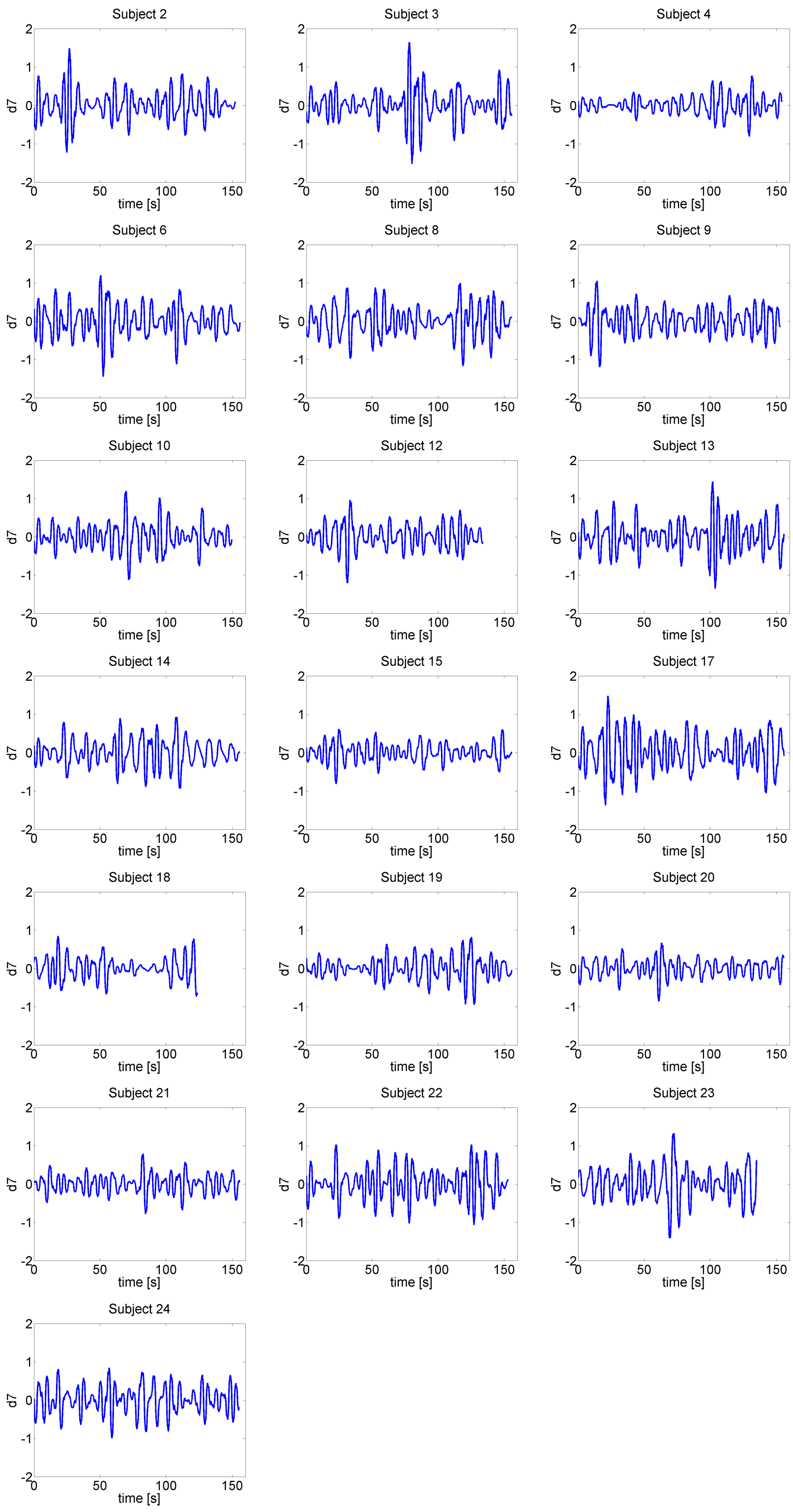

Figure 4.9: Fatigue waves (d7) for each participating subject over time in seconds. Subjects 1, 5, 11 and 16 did not participate in this part of the study and subject 7 was excluded because of invalid data.

As mentioned in chapter *3.1.3 Evaluation Methods*, d10 is only the "best guess" for detail level. Figure 4.8 and Figure 4.9 show, that detail levels d5 and d7 also indicate fatigue waves that match the definition (see chapter *2.2.2 The Pupil and Pupillary Oscillations*). d5 showed fatigue waves of about 0.5 to 1 seconds, d7 showed fatigue waves of about 4 seconds. So using d10 as "best guess" seemed to be an arbitrary decision and could, for that reason, not be considered dependable.

So fatigue waves could not be seen as a well-performing indicator for fatigue within this experiment. Only a10 could be taken as an indicator for cognitive workload within visual exploration tasks as it showed a decrease in the average values of the pupil diameter. Also the average value is stable with different detail levels, so the choice of the wavelet decomposition level is not that relevant.

4.1.3 Horizontal Distance between Left and Right Eye (Vergence Accuracy)

The horizontal distance between left and right eye illustrates the vergence accuracy over time and is shown in Figure 4.10. At first glance, this parameter seemed constant over time, but in the end of the recording there were more peaks (for instance a large one at about 105 seconds). A *linear* trend line was applied. It is not known, if a *linear* relation actually appeared. Thus, a trend was shown but no data modelling was performed. The trend line had the equation

$$vac = 0.00085\ t - 0.05 \tag{4.3}$$

$vac = normed\ vergence\ accuracy\ [nondimensional]$
$t = time\ [s]$

and therefore showed an increase of the vergence accuracy between left and right eye. This was confirmed by violin plots (see Figure 4.11) for comparison of first and last tenth of the recording time. There was no significant difference between these segments ($p = 0.0246$, ANOVA), although a difference of $\Delta = 0.0985$ for the median of first and last tenth of the recording time appeared (median of the first tenth: $med_1 = -0.03492$, median of the last tenth: $med_2 = 0.06358$). For that, a relevant trend towards an increasing divergence was shown. It was assumed that this could come from increasing phoria of the subjects due to increasing fatigue but that could not be proven.

Table 4.1: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 3. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.0431	-0.0286	vac	1.0000	-0.0444	-0.0331
gp-x	0.0431	1.0000	0.4720	gp-x	-0.0444	1.0000	0.4017
gp-y	-0.0286	0.4720	1.0000	gp-y	-0.0331	0.4017	1.0000

However, the variance in the signal was high. The effect could only be observed in the combined data of all subjects, not on an individual level. Furthermore, the effect size was quite small, what made the usability of this parameter questionable.

The Spearman correlation for each subject was calculated.
The results of subject 3 are described in Table 4.1 as an example. For the complete results of all subjects see appendix *A Results of Correlation Analysis for Vergence Accuracy and Gaze Position* for reasons of a clear arrangement.
Correlation coefficients appeared in absolute values from $R = 0.0286$ to $R = 0.4720$ in this subject and showed a high variance for subject 3 as for all other subjects. Significance levels were not calculated, as they would only refer to one individual subject and one individual direction, respectively. This is not relevant within this study as there is no prediction for the combined data of all subjects for the reason of inter-individual variations.

Correlation analysis between vergence accuracy and gaze position suggested correlations, but neither for intra-individual nor for inter-individual comparison on a consistent basis. Variance in correlation coefficients was high. This again suggested an insufficient determination of the actual vergence angle.

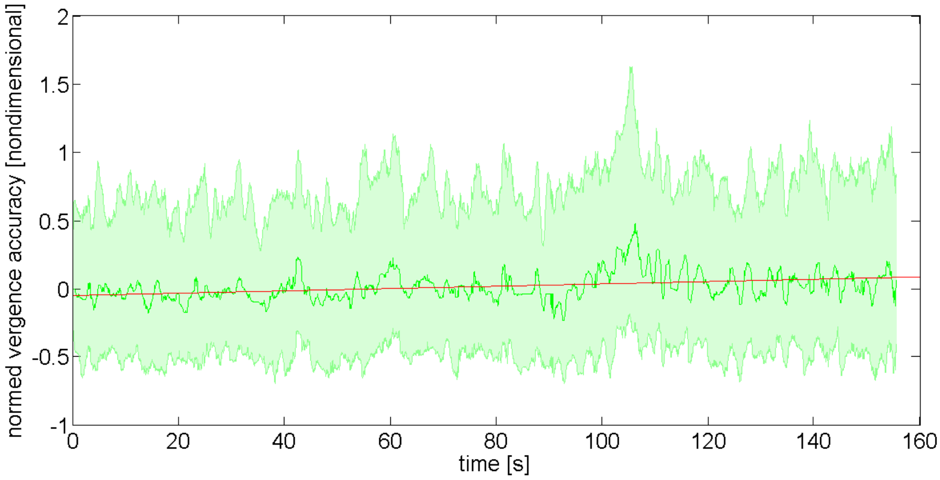

Figure 4.10: Median (50th percentile, dark green) over all subjects as well as 20th and 80th percentiles (light green) for normed horizontal distance between left and right eye (vergence accuracy, nondimensional) over time in seconds with trend line (red). Vergence accuracy was normed by subtraction of its median and division by its standard deviation afterwards (see chapter *3.1.3 Evaluation Methods*).

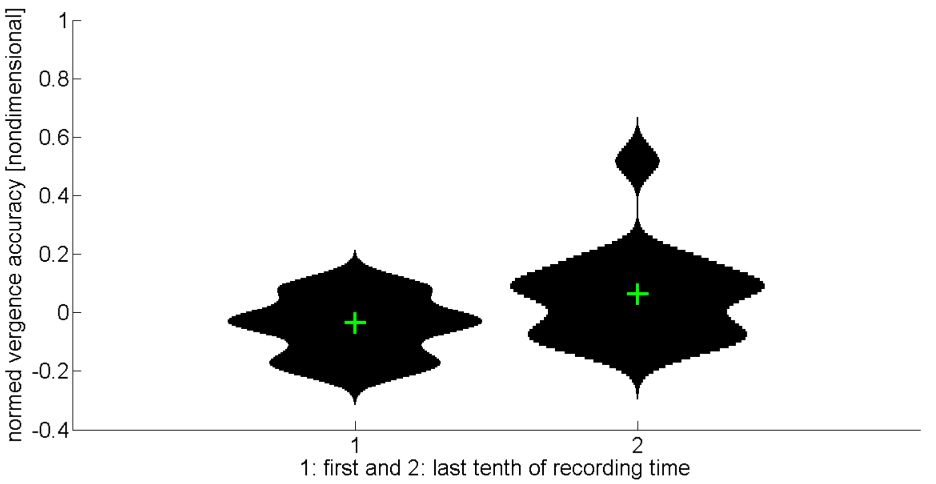

Figure 4.11: Violin plots (as explained in Figure 4.3) for comparison of first [0 s;15 s] and last [140 s;155 s] tenth of recording time concerning normed horizontal distance between left and right eye (vergence accuracy, nondimensional, medians are represented by green crosses). Vergence accuracy was normed by subtraction of its median and division by its standard deviation afterwards (see chapter *3.1.3 Evaluation Methods*).

So this parameter, contrary to expectations, could neither be taken as a parameter for vigilance nor for attention for eye tracking measurement, nor for individual notification of fatigue.

The small peaks appearing for the median of vergence accuracy probably make sense as we expected this distance to be larger at the beginning and end of saccades.

4.1.4 Blink Rate

The blink rate appeared as steps because of the way it was calculated. When using a sliding window of 15 seconds and upscaling to blinks per minute, the count of blinks during the 15 seconds time interval was multiplied by four. Figure 4.12 illustrates the median over all subjects as well as 20^{th} and 80^{th} percentiles for the blink rate. Median blink rate was approximately eight blinks per minute at the beginning of the session. Towards the end of the recording session, there was a higher variation of the median blink rate.

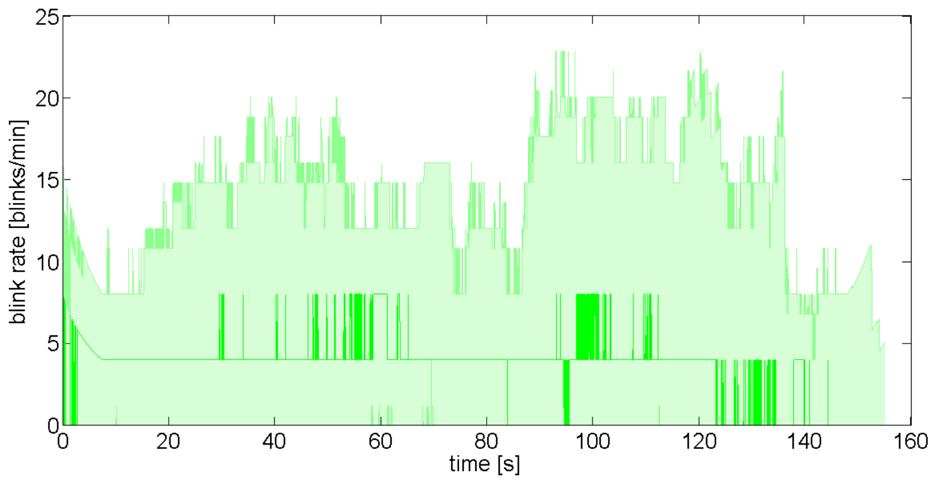

Figure 4.12: Median (50^{th} percentile, dark green) over all subjects as well as 20^{th} and 80^{th} percentiles (light green) for blink rate in blinks per minute over time in seconds

There was no significant difference between first and the last tenth of recording time, what can be seen in Figure 4.13 ($p = 0.8443$, ANOVA). In addition, no difference between the median of first and last tenth of the recording time appeared (median of the first and last tenth: $med_{1,2} = 0\ blinks/min$. For that, no relevant trend was shown.

The blink rate was generally quite low with under 10 blinks per minute by median but showed

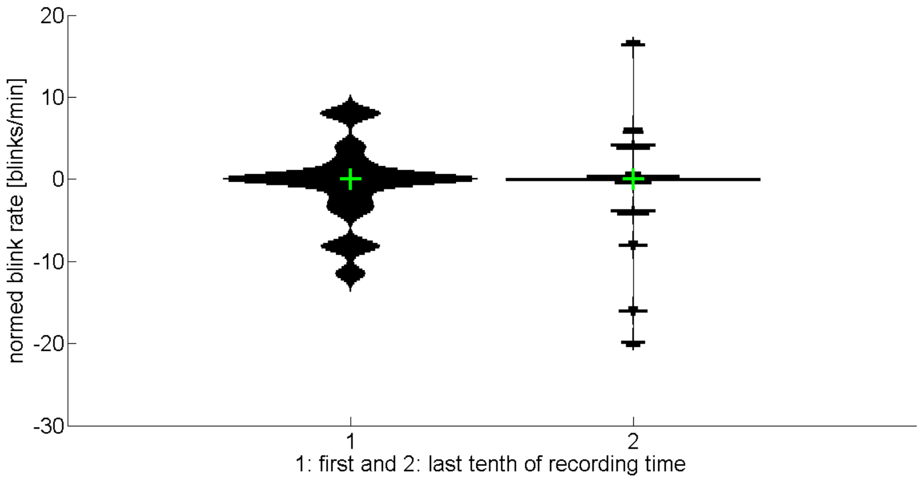

Figure 4.13: Violin plots (as explained in Figure 4.3) for comparison of first [0 s;15 s] and last [140 s;155 s] tenth of recording time concerning normed blink rate in blinks per minute. Medians are represented by green crosses. Blink rate was normed by subtraction of its median (see chapter *3.1.3 Evaluation Methods*).

an increase towards the end of the recording time. Yet there were high variations, as the enlargement of the 20th and 80th percentiles demonstrated. But there was no relevant increase of the blink rate median, so only some of the subjects showed an increasing blink rate.

According to the results of this experiment, blink rate could not be seen as a well-performing indicator for vigilance and fatigue for each subject. For that, the evaluation of fatigue by means of the blink rate could be a feasible method but not necessarily for image viewing tasks and in general not without fail.

4.1.5 Blink Duration

Figure 4.14 shows the median as well as the 20th and 80th percentiles for the blink duration. At first glance, this parameter seemed constant over time, except the last 30 seconds: After about 120 seconds of the recording time, a distinct decrease appeared. This was not explainable by implication. The observed effect only appeared when combining the data of all subjects, not for the individual.
After about 150 seconds, a distinct decrease in blink duration seemed to appear. That was due to the usage of a sliding window for the calculation of this parameter. For that, this can be seen as an artefact.

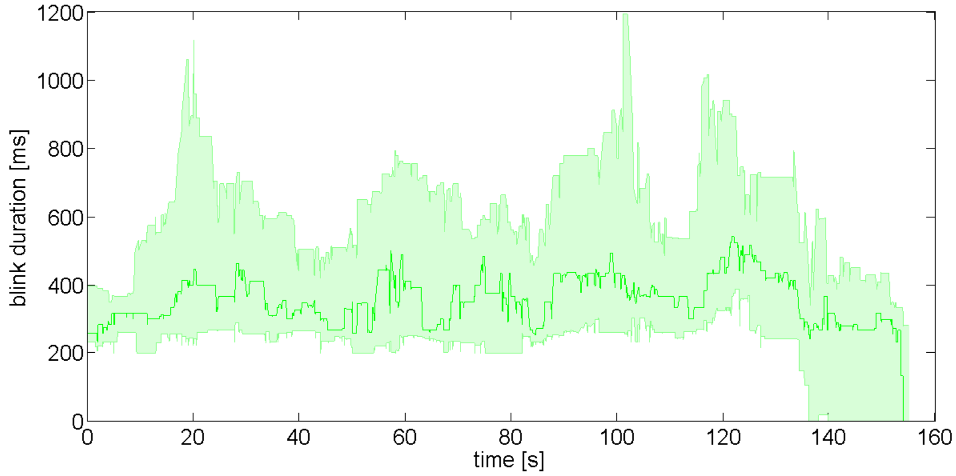

Figure 4.14: Median (50th percentile, dark green) over all subjects as well as 20th and 80th percentiles (light green) for blink duration in milliseconds over time in seconds

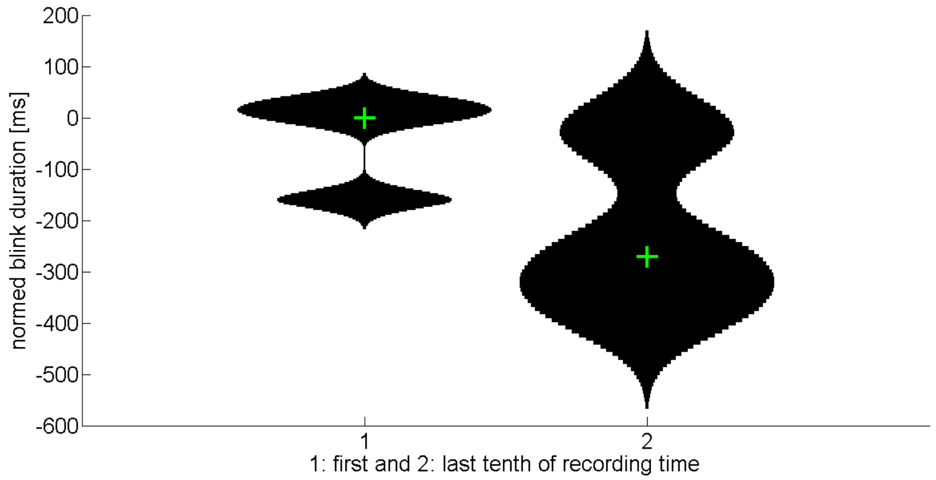

Figure 4.15: Violin plots (as explained in Figure 4.3) for comparison of first [0 s;15 s] and last [140 s;155 s] tenth of recording time concerning normed blink duration in milliseconds. Medians are represented by green crosses. Blink duration was normed by subtraction of its median (see chapter *3.1.3 Evaluation Methods*).

First tenth was compared to last tenth of the recording time. Violin plots (see Figure 4.15) show, that there was no difference between first and the last tenth of the recording time ($p = 0.2033$, ANOVA): Blink duration was not significantly shorter but also more variant in the beginning of the recording session, although a difference of $\Delta = -270.4\ ms$ for the median of first and last tenth of the recording time appeared (median of the first tenth: $med_1 = 0\ ms$, median of the last tenth: $med_2 = -270.4\ ms$). For that, a relevant trend was shown.

So blink duration could be seen as a well-performing parameter showing trends for attention within experimental setups for visual exploration examinations for many subjects.

4.1.6 Saccade Length

Figure 4.16 illustrates the median over all subjects as well as 20th and 80th percentiles for the saccade length averaged over all subjects. In the beginning of the eye tracking sessions, more peaks were detected and the median saccade length seemed to be longer than at the end of the recording session. A *linear* trend line was applied. It is not known, if a *linear* relation actually appeared. Thus, a trend was shown but no data modelling was performed. A slightly decreasing saccade length appeared as the equation of the trend line was

$$sl = -0.53\ t + 230 \tag{4.4}$$

$sl = saccade\ length\ [px]$
$t = time\ [s]$

To compare first and last tenth of the recording time, violin plots were created and are shown in Figure 4.17. These violin plots showed no significant difference between first and last tenth of recording time ($p = 0.0112$, ANOVA). Yet a difference of $\Delta = -0.3135\ px$ for the median of first and last tenth of the recording time appeared (median of the first tenth: $med_1 = 0.1898\ px$, median of the last tenth: $med_2 = -0.1237\ px$) and confirmed the decrease over time the applied trend line indicated. For that, a relevant trend was shown.

Median of saccade length was not significantly shorter during the last tenth of the recording, although a trend appeared.
As the saccade length decreased over time and was, by trend, longer during the first than during the last tenth of the recording time, it was assumed that in the beginning of a session,

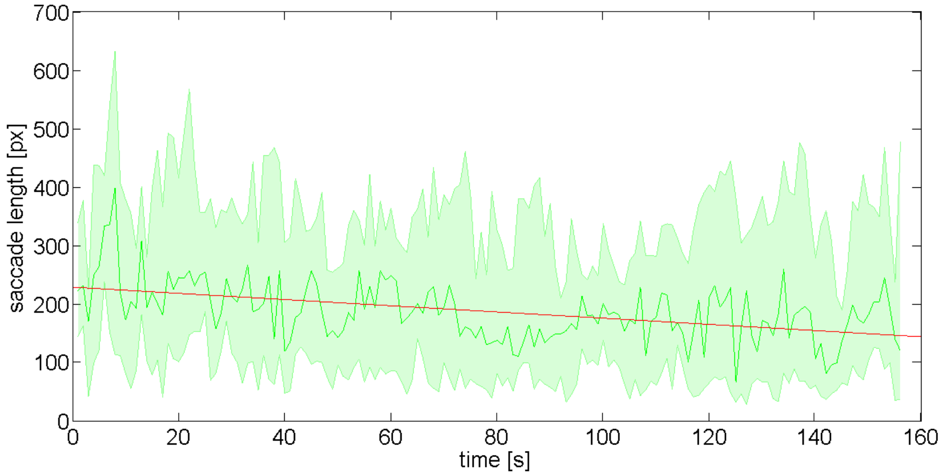

Figure 4.16: Median (50th percentile, dark green) over all subjects as well as 20th and 80th percentiles (light green) for saccade length in pixels over time in seconds with trend line (red)

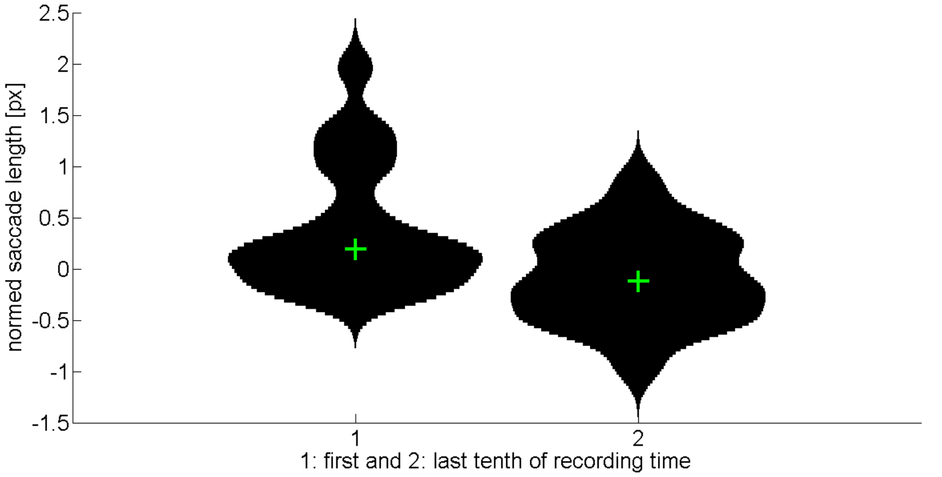

Figure 4.17: Violin plots (as explained in Figure 4.3) for comparison of first [0 s;15 s] and last [140 s;155 s] tenth of recording time concerning normed saccade length in pixels. Medians are represented by green crosses. Saccade length was normed by subtraction of its median (see chapter *3.1.3 Evaluation Methods*).

subjects tried to get an overview over the picture presented by long saccades and later, towards the end of the recording sessions, they got into details, what required shorter saccades. This was yet observable for the median over all subjects, but not for the individual. Therefore, the saccade length can be seen as a useful parameter for attention within experimental setups for many subjects. For individual predictions of attention, the parameter can be used, but not without fail.

4.1.7 Fixation Duration

Figure 4.18 shows the fixation duration (raw data and average smoothed) for each subject over time. Variations for each subject occured, but there were subjects with lower (e.g. subjects 2, 6, 8, 9, 14 and 17) and higher duration peaks (e.g. subjects 3, 10, 12 and 18). By trend, most of the peaks occured towards the end of the recording sessions.

Figure 4.19 shows the median of fixation duration of all subjects over time. Fixation duration increased, and peaks in durations of over 1.5 seconds in median appeared after 140 seconds of the recording sessions. In addition, there was a distinct enlargement of the distance between 20^{th} and 80^{th} percentile after about 115 seconds. The 80^{th} percentile showed values exceeding five seconds for fixation duration at about 115 seconds.
First and last tenth of the recording time were compared, violin plots were created (see Figure 4.20). They showed quite similar median values but one outlier of over 20 seconds for the last tenth of the recording time. No significant effect ($p = 0.0197$, ANOVA) occured for the comparison of these segments of fixation duration, although a difference of $\Delta = 186.5\ ms$ for the median of first and last tenth of the recording time appeared (median of the first tenth: $med_1 = 410.8\ ms$, median of the last tenth: $med_2 = 597.3\ ms$). For that, a relevant trend was shown.
The trend showed that subjects began to stare as they got bored, and for that reason, unconcentrated and inattentive. After a certain time, which was obviously 115 seconds in this experiment, subjects began to deal with their lack of concentration in different ways: Some tried to re-concentrate on their task, some did not. The latter ones just stared, not anxious to miss more cognitive useful information. An example for that is an outlier detected that showed a fixation duration of over 20 seconds what obviously was staring.

Fixation duration could, according to this study, be seen as a well-performing parameter showing trends for attention for a group of several subjects as well as for the individual within eye tracking experiments.

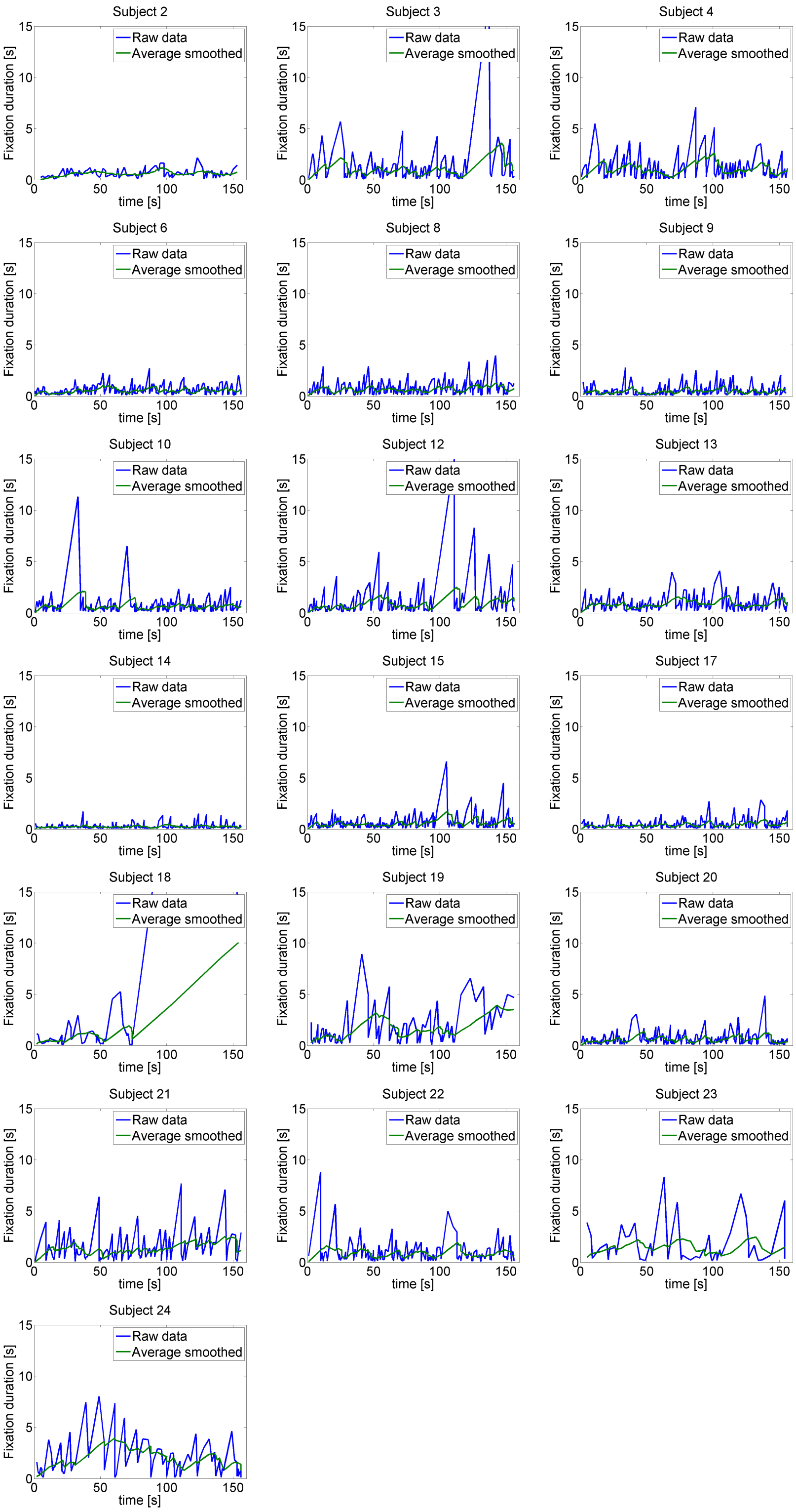

Figure 4.18: Fixation duration in ms (blue: raw data and green: average smoothed) for each participating subject over time in seconds. Subjects 1, 5, 11 and 16 did not participate in this part of the study and subject 7 was excluded because of invalid data.

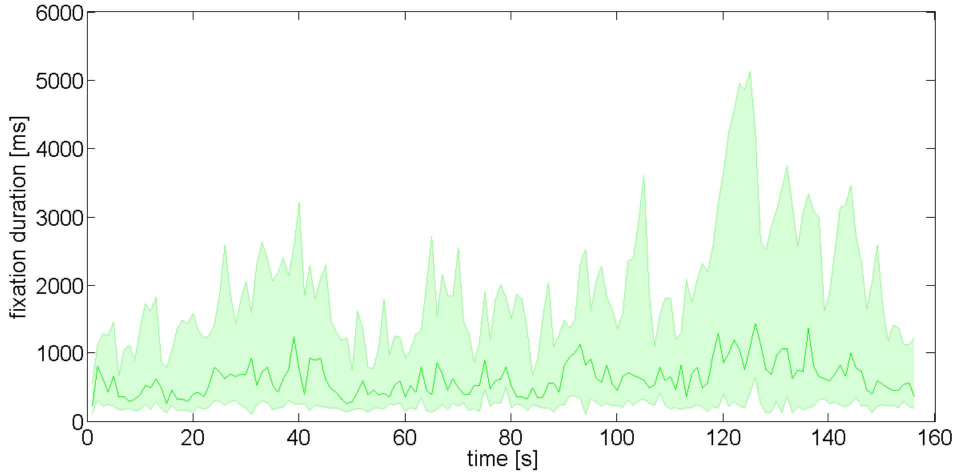

Figure 4.19: Median (50th percentile, dark green) over all subjects as well as 20th and 80th percentiles (light green) for fixation duration in milliseconds over time in seconds

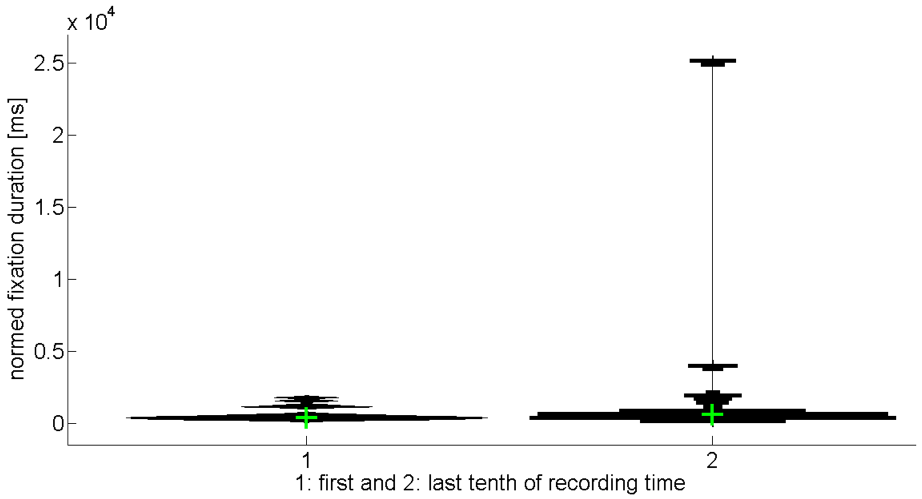

Figure 4.20: Violin plots (as explained in Figure 4.3) for comparison of first [0 s;15 s] and last [140 s;155 s] tenth of recording time concerning normed fixation duration in milliseconds. Medians are represented by green crosses. Fixation duration was normed by subtraction of its median (see chapter *3.1.3 Evaluation Methods*).

4.1.8 Index of Cognitive Activity (ICA)

Figure 4.21 shows the median over all subjects as well as 20th and 80th percentiles for the Index of Cognitive Activity (ICA). The ICA decreased after about 85 out of 155 seconds altogether, which makes approximately half of the recording time, and re-increased after about 120 seconds. A *linear* trend line was applied. It is not known, if a *linear* relation actually appeared. Thus, a trend was shown but no data modelling was performed. A slightly decreasing ICA appeared as the equation of the trend line was

$$ICA = -0.0055\, t + 8.4 \tag{4.5}$$

$ICA = Index\ of\ Cognitive\ Activity\ [nondimensional]$
$t = time\ [s]$

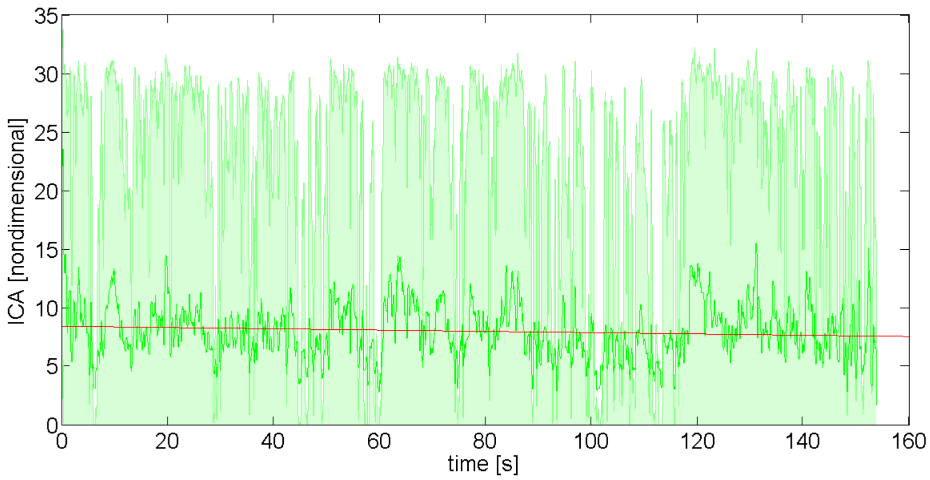

Figure 4.21: Median (50th percentile, dark green) over all subjects as well as 20th and 80th percentiles (light green) for Index of Cognitice Activity (nondimensional) over time in seconds with trend line (red)

First tenth of the recording time was tested against the last tenth. There is no significant defference between these segments concerning the Index of Cognitive Activity (ICA) at all ($p = 0.5785$, ANOVA, see also Figure 4.22), and also no difference for the median of first and last tenth of the recording time appeared (median of the first and last tenth: $med_{1,2} = 0$). For that, no relevant trend was shown.

So, more cognitive workload in the beginning of the recording time for this task was assumed.

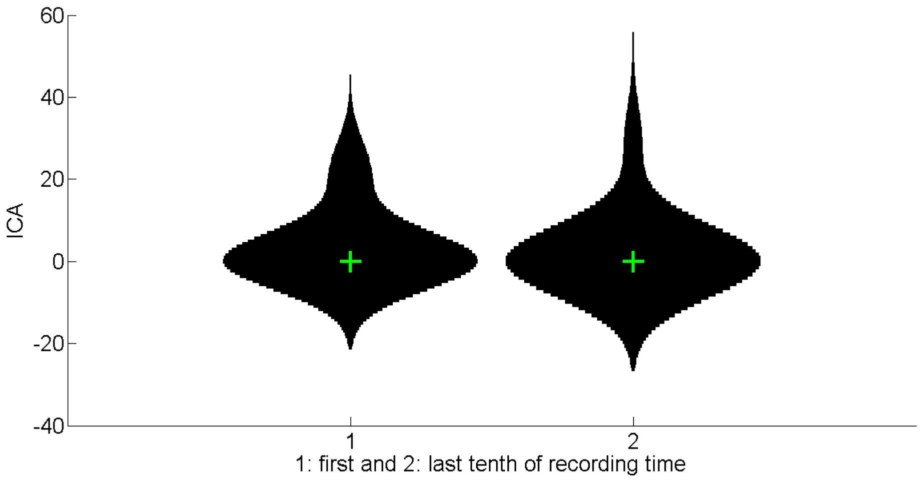

Figure 4.22: Violin plots (as explained in Figure 4.3) for comparison of first [0 s;15 s] and last [140 s;155 s] tenth of recording time concerning Index of Cognitice Activity (nondimensional, medians are represented by green crosses)

That was illustrated by a workload peak in the beginning. However, these suggestions could not be seen as valid as no significant differences between first and last tenth of the recording timeand no relevant trend appeared. This was probably caused by level and shape of the implemented wavelet in the ICA. Level and shape are due to the frame rate and optimized for another frame rate than ours.

So the Index of Cognitive Activity (ICA) could not be seen as a well-performing parameter for the median over various subjects, and in addition predictions concerning cognitive workload could not be made for the individual.

4.2 Re-analyzing Data: Perimetry

As shown, there are parameters that perform well as indicators for vigilance, attention and cognitive workload. Not all of them could be applied to perimetry data because for the detection of some parameters, such as vergence, *binocular* eye tracking was needed. The camera currently installed in the OCTOPUS 900 perimeter can only record one eye, additionally perimetry is most frequently recorded *monocularly*. Therefore with the current setup we could only investigate the measures that can be derived from one eye. Furthermore measures that require eye movement are not suited for a perimetric examination, since a constant central fixation is required. So saccade length and fixation duration could not be taken into account as indicators for fatigue during a perimetric examination.

The parameters mentioned in chapter *3.2.3 Evaluation Methods* were applied.
In contrast to the previous experiment there was another parameter available: the error rate (false positive and false negative catch trials) during the examination could be utilized as a measure of the subjects' performance. This performance measure could be correlated with attention and vigilance indicators.

Subsequently, the results concerning the particular parameters are shown for each subject examined. First, pupillary-based, then blink-based parameters are shown.

4.2.1 Fatigue Waves (a10, d10), Pupil Diameter Variability and Their Correlation to Error Rate

Subject 1

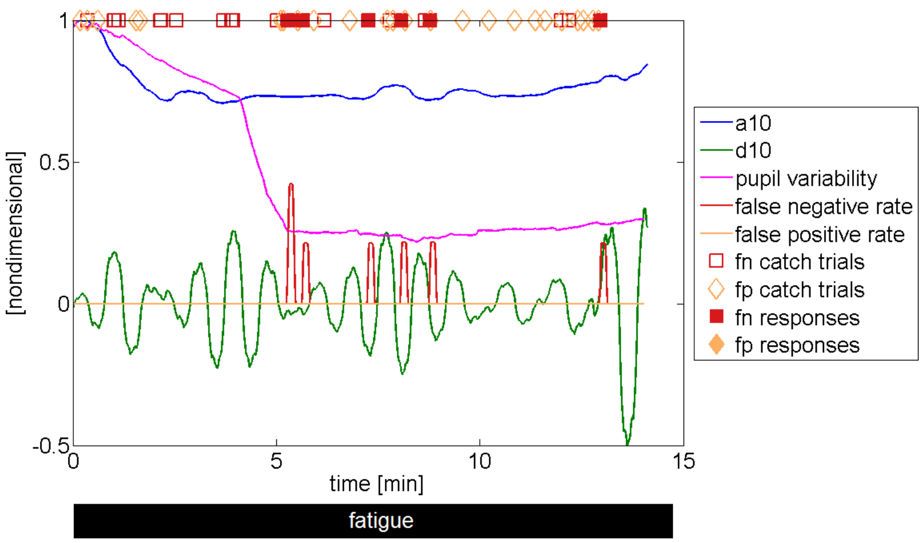

Figure 4.23: Normed fatigue waves a10 (blue) and d10 (green), pupil diameter variability (pink, all nondimensional) and error rates (fn: red and fp: orange, nondimensional) over time in minutes for subject 1, time period of assumed fatigue marked by black bar, y-axis scaled for all parameters noticed on the right. Pupil diameter variability, error rates and a10 were normed by their maximum, d10 was normed by twice its maximum.

Figure 4.23 shows fatigue waves and pupil diameter variability as well as the error rate (separated into false negative and false positive errors) for subject 1. In the beginning of the recording session, pupil diameter variability was high. Detail level d10 showed fatigue waves during the whole recording session. The values for a10 decreased. Many false negative responses to catch trials (26 %) appeared throughout the whole recording time, whereas there were no false positive responses to catch trials. For that, the number of false positive responses was acceptable, whereas the number of false negative responses was on an inacceptable level.

In addition, the error rates are shown and were correlated to a10 and the pupil diameter variability.
For a10 and the error rate, the calculated correlation coefficient was $R = -0.1013$. For pupil diameter variability and error rate, the calculated coefficient was $R = -0.2771$.

Subject 2

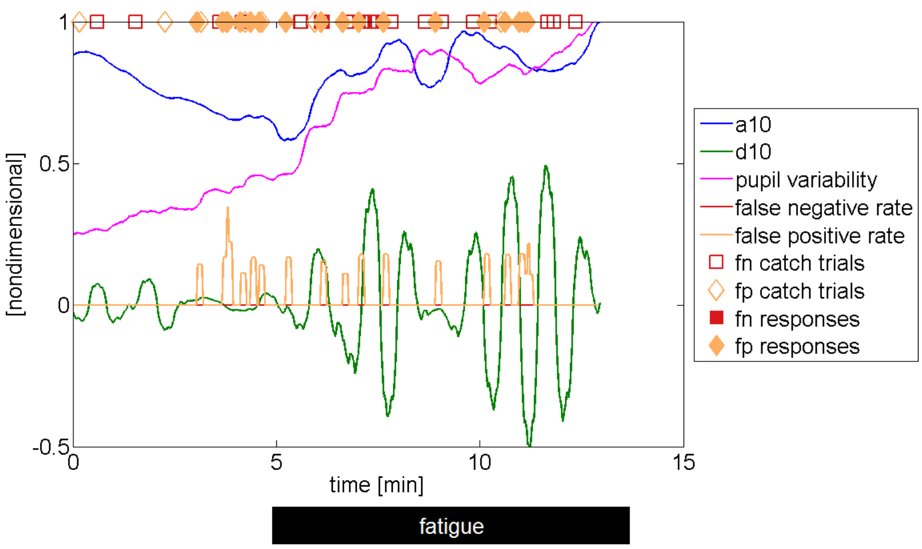

Figure 4.24: Normed fatigue waves a10 (blue) and d10 (green), pupil diameter variability (pink, all nondimensional) and error rates (fn: red and fp: orange, nondimensional) over time in minutes for subject 2, time period of assumed fatigue marked by black bar, y-axis scaled for all parameters noticed on the right. Pupil diameter variability, error rates and a10 were normed by their maximum, d10 was normed by twice its maximum.

Figure 4.24 shows fatigue waves and pupil diameter variability as well as the error rates for subject 2. Pupil diameter variabiliy showed higher values towards the end of the recording session. Fatigue waves (d10) appeared after about six minutes, subjectively rated. The values for a10 first decreased, then showed long-term oscillations. No false negative responses to catch trials appeared during the whole recording session (0 %). False positive responses to catch trials appeared cumulative with the occurrence of fatigue after about six minutes. 55 % false positive responses to catch trials from the beginning to six minutes and 69 % false positive responses to catch trials from six minutes on were given. So an increase by 14 percentage points appeared. However, both false positive values are on an inacceptable level and for that the increase in false positive rate could not be seen as relevant.

For a10 and error rate, the calculated correlation coefficient was $R = -0.1475$. For pupil diameter variability and error rate, the calculated coefficient was $R = 0.0431$.

Subject 3

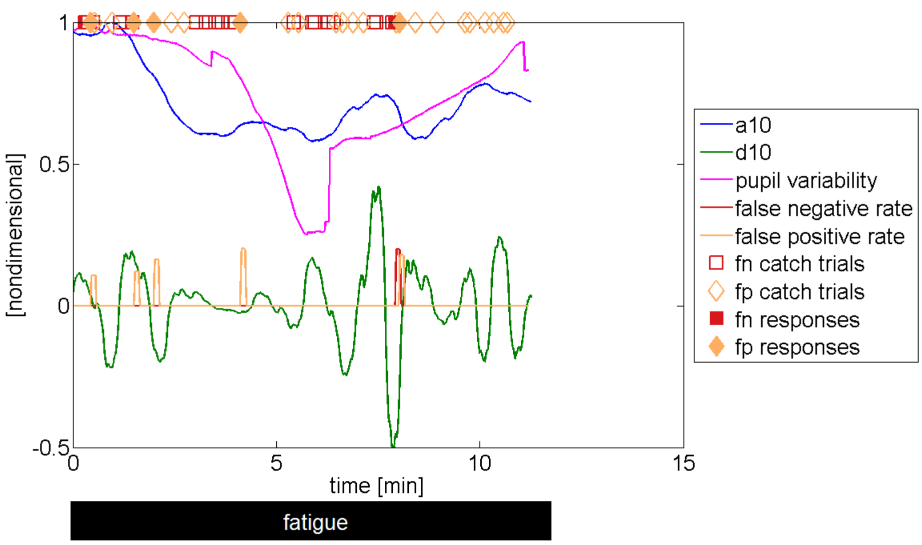

Figure 4.25: Normed fatigue waves a10 (blue) and d10 (green), pupil diameter variability (pink, all nondimensional) and error rates (fn: red and fp: orange, nondimensional) over time in minutes for subject 3, time period of assumed fatigue marked by black bar, y-axis scaled for all parameters noticed on the right. Pupil diameter variability, error rates and a10 were normed by their maximum, d10 was normed by twice its maximum.

Figure 4.25 shows fatigue waves and pupil diameter variability as well as the error rates for subject 3. Pupil diameter variability showed higher values in at the beginning and towards the end of the recording time. Fatigue waves (d10) appeared during the whole perimetric session. The values for a10 first decreased, then showed long-term oscillations. False negative responses to catch trials (7 %) appeared as well as false positive responses to catch trials (19 %) throughout the whole recording session. But both false positive and false negative rate were on an acceptable level.

For a10 and error rate, the calculated correlation coefficient was $R = 0.1126$. For pupil diameter variability and error rate, the calculated coefficient was $R = 0.1276$.

Subject 4

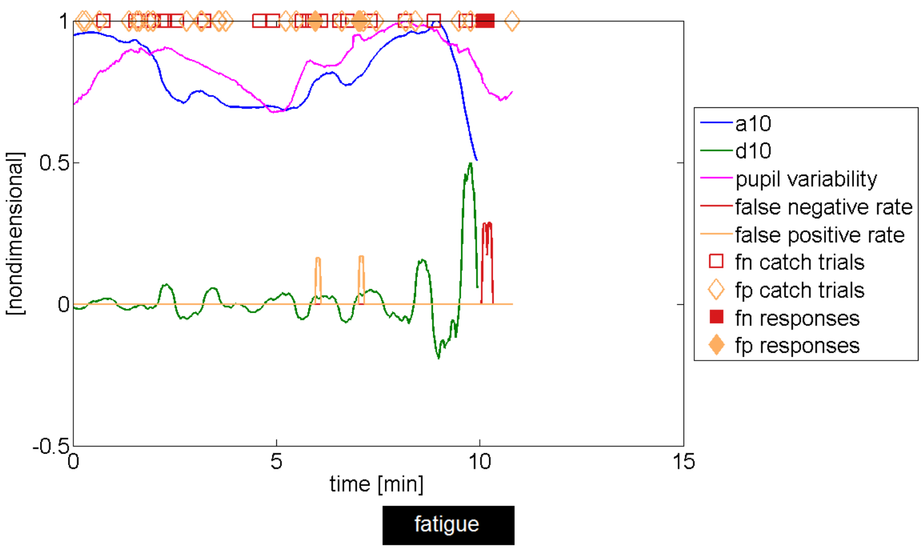

Figure 4.26: Normed fatigue waves a10 (blue) and d10 (green), pupil diameter variability (pink, all nondimensional) and error rates (fn: red and fp: orange, nondimensional) over time in minutes for subject 4, time period of assumed fatigue marked by black bar, y-axis scaled for all parameters noticed on the right. Pupil diameter variability, error rates and a10 were normed by their maximum, d10 was normed by twice its maximum.

Figure 4.26 shows fatigue waves and pupil diameter variability as well as the error rates for subject 4. Pupil diameter variability showed a peak after about eight minutes of the recording time. Fatigue waves (d10) also began to appear after about eight minutes, subjectively rated. The values for a10 first decreased, then re-increased towards the end of the perimetric session. From the beginning of the recording until the appearance of fatigue waves, no false negative and 9 % false positive responses to catch trials were given, which was acceptable. From the appearance of fatigue waves after about eight minutes, 50 % false negative and no false positive responses to catch trials were given. So the number of false positive catch trials was still acceptable and the false negative rate increased to an inacceptable level.

For a10 and error rate, the calculated correlation coefficient was $R = -0.0276$. For pupil diameter variability and error rate, the calculated coefficient was $R = -0.0630$.

Subject 5

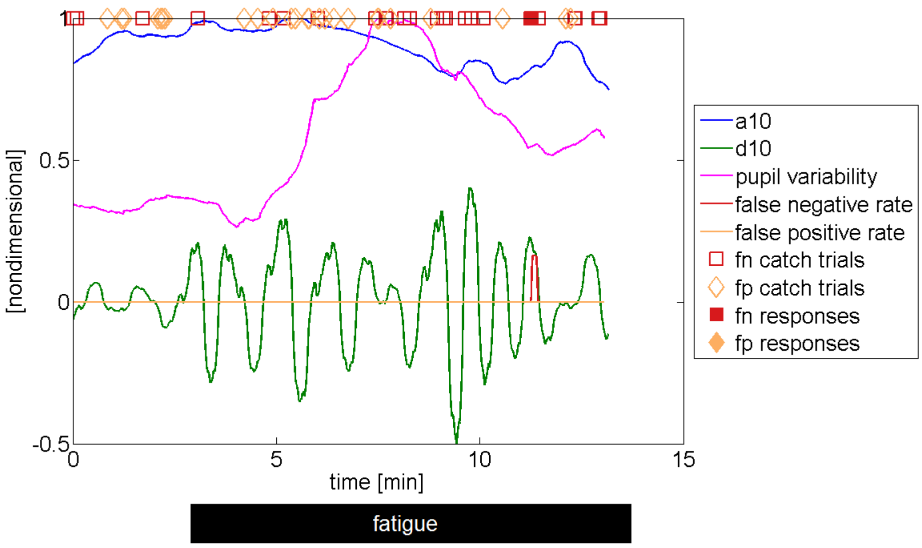

Figure 4.27: Normed fatigue waves a10 (blue) and d10 (green), pupil diameter variability (pink, all nondimensional) and error rates (fn: red and fp: orange, nondimensional) over time in minutes for subject 5, time period of assumed fatigue marked by black bar, y-axis scaled for all parameters noticed on the right. Pupil diameter variability, error rates and a10 were normed by their maximum, d10 was normed by twice its maximum.

Figure 4.27 shows fatigue waves and pupil diameter variablitiy as well as the error rates for subject 5. Pupil diameter variability increased after five minutes and decreased towards the end of the recording session. Fatigue waves (d10) appeared after about three minutes, subjectively rated. The values for a10 were at a quite constant level and showed, for that, no relevant variability. From the beginning of the recording until the appearance of fatigue waves, 0 % false negative responses to catch trials were given, which was acceptable. From the appearance of fatigue waves after about three minutes, 4 % false negative to catch trials were given. So the number of false negative catch trials was still acceptable. During the whole perimetric session no false positive responses to catch trials appeared (0 %), what was highly acceptable.

For a10 and error rate, the calculated correlation coefficient was $R = -0.1244$. For pupil diameter variability and error rate, the calculated coefficient was $R = 0.0068$.

Subject 6

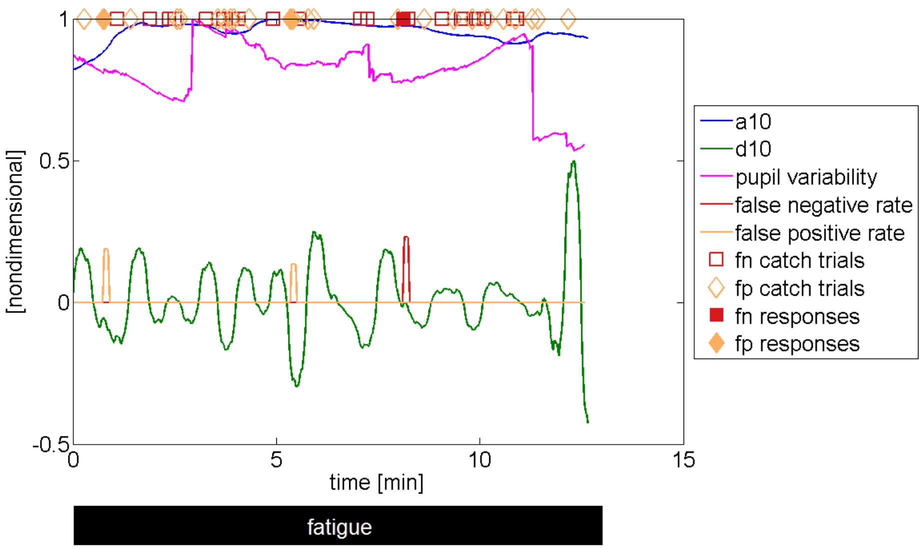

Figure 4.28: Normed fatigue waves a10 (blue) and d10 (green), pupil diameter variability (pink, all nondimensional) and error rates (fn: red and fp: orange, nondimensional) over time in minutes for subject 6, time period of assumed fatigue marked by black bar, y-axis scaled for all parameters noticed on the right. Pupil diameter variability, error rates and a10 were normed by their maximum, d10 was normed by twice its maximum.

Figure 4.28 shows fatigue waves and pupil diameter variability as well as the error rates for subject 6. Pupil diameter variability comparatively high throughout the whole recording session. Fatigue waves (d10) appeared from the beginning to the end. The values for a10 were at a quite constant level and showed, for that, no relevant variability. False negative responses to catch trials (4 %) appeared as well as false positive responses to catch trials (7 %) throughout the whole recording session. But both false positive and false negative rate were on an acceptable level.

For a10 and error rate, the calculated correlation coefficient was $R = 0.0205$. For pupil diameter variability and error rate, the calculated coefficient was $R = -0.0724$. For that, there was a low positive correlation.

Summary

By trend, pupil diameter variability and fatigue waves (d10) appeared together as they were both calculated from changes in the pupil diameter and its change and were therefore related to each other.

False negative and false positive responses to catch trials can hardly be related to fatigue waves (d10). For both false positive and false negative responses to catch trials, no dependable result can be stated, as fatigue waves over a limited time span could be observed in only three out of six subjects. False positive and false negative rates were expected to increase with increasing fatigue. As values up to 20 % for false responses were rated as acceptable, values larger than 30 % were rated as inacceptable, values were supposed to be lower than 20 % when no fatigue occured and higher than 30 % when fatigue occured. An increase of false positive responses to catch trials (14 percentage points) after the occurrence of fatigue wave could only be observed in one subject, yet the percentage of false positive responses was inacceptable before fatigue occured. For that, this increase in false positive responses was not relevant. An increase in false negative responses to catch trials could also be observed in one subject (50 percentage points). In this subject, the false negative rate was acceptable befpre fatigue occured and inacceptable afterwards. Only one subject who showed fatigue waves throughout the whole perimetric session showed an inacceptably high false negative rate. All other subjects showing fatigue from the beginning to the end of the perimetric examination showed acceptable false positive and false negative rates.

The average values (a10) decreased in most of the cases as a result of adaption to a brighter environment. Subject 4 showed a re-increasing average value (a10) and subjects 2 and 3 showed long-term oscillations. That was not explainable by implication.
For a10 and error rate correlation analysis showed negative coefficients in the majority of subjects. For that, a10 and error rate seemed to be related during a perimetric session: With a decrease of a10, error rate increased. But the correlation was on a low and inter-individual varying level and for that could also be due to the fact, that a10 decreased with the recording time for reasons of adaption to the background illuminance level of the perimeter cupula. Error rate increased as fatigue increased over time. For that, no reliable statement could be made for the relation between error rate and a10 with respect to vigilance.

Significance levels for all correlations have not been calculated, as they are not relevant. Only significance on an intra-individual basis would have been shown, not for the combined data of all subjects. Due to the small sample size and large inter-individual differences, only a descriptive statistical consideration was performed at this juncture.

4.2.2 Blink Rate and its Correlation to Error Rate

Blink rate was calculated as seen in chapter *3.2.3 Evaluation Methods* and is shown in Figure 4.29 (absolute values, not normed). For subject 1, an increasing blink rate over time with peaks at 15 blinks/min appeared. For subjects 2, 3, 4, 5 and 6, very low blink rates of 0 to 3 blinks per minute appeared. There were even time slots where no blinks appeared for more than 2 minutes, what is possible but astounding. Perimetry data taken from Müller et al. 2014 actually showed no invalid data points that are defined as blinks during these time slots. Possibly, the algorithm used for the recording of perimetry data was not able to detect blinks correctly in every case.

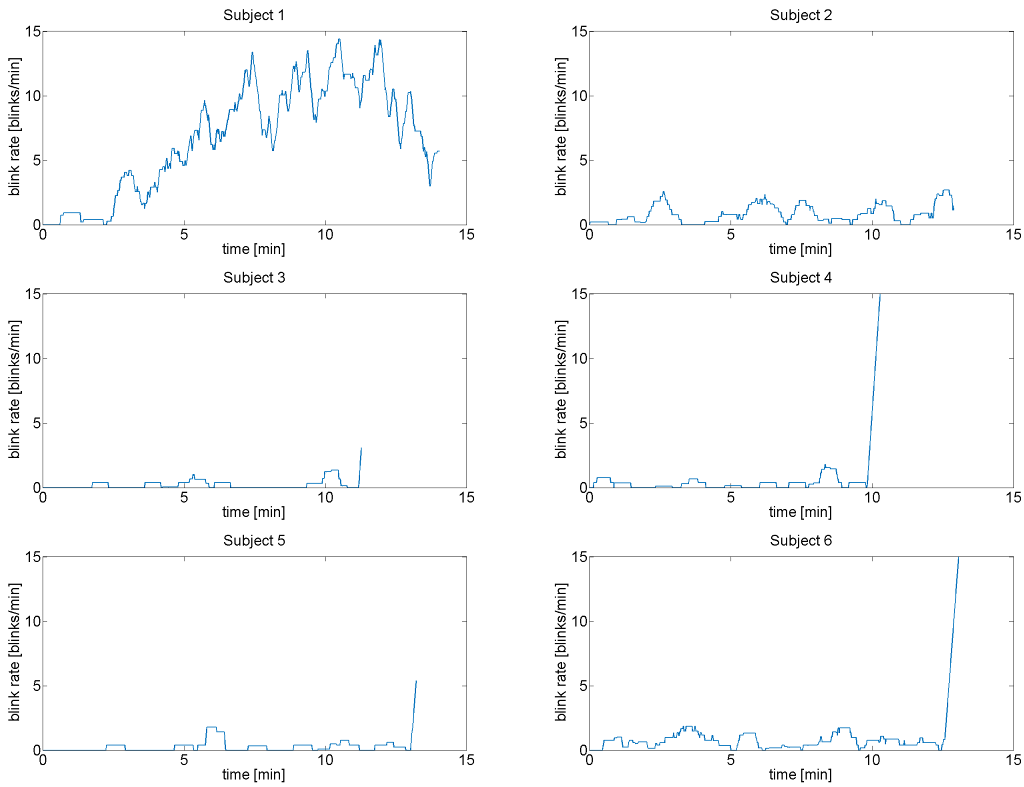

Figure 4.29: Absolute values of blink rate in blinks per minute over time in minutes for each subject

Subject 1

Figure 4.30 shows error rates and blink rate for subject 1. The calculated correlation coefficient was $R = 0.2712$. An increase of the error rate with increasing blink rate, and for that a positive correlation, was assumed and actually appeared. However, a more distinct correlation and for that a higher correlation coefficient was expected. So there has to be a reason for these results: Subject 1 might have blinked after every stimulus perceived, even when getting bored and tired. To visualize that, Figure 4.31 shows the detection rate (a normed value for the number of stimuli perceived) and the blink rate for subject 1. Peaks of the blink rate seemed to occur together with peaks of the detection rate, for instance after about three, six and ten minutes. Spearman correlation analysis for blink rate and detection rate showed a correlation coefficient of $R = 0.85$ and for that a high correlation. For visualization see Figure 4.32. Measurement values seemed to appear in steps as the blink rate was calculated using a sliding window technique. This correlation was also significant with a p-value of $p = 0.0000$. So subject 1 may have blinked after every stimulus perceived, even when getting tired.

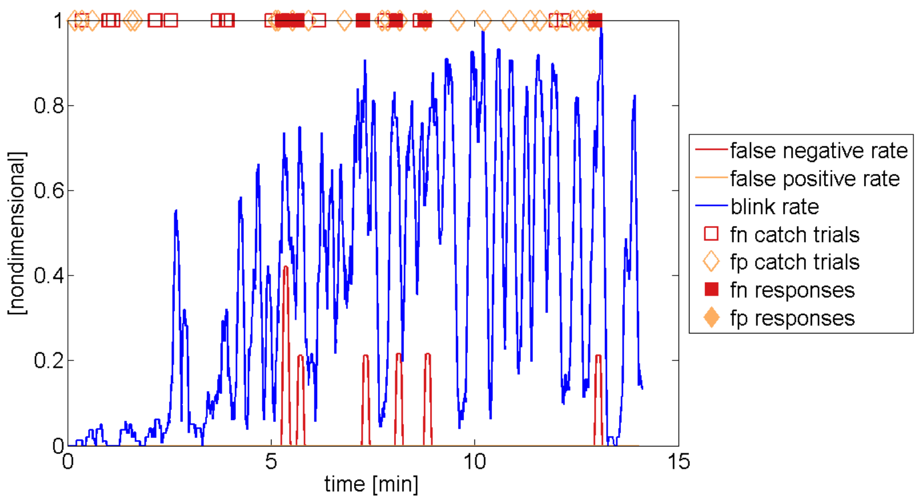

Figure 4.30: Normed error rates (nondimensional, fn: red and fp: orange) and blink rate (nondimensional, blue) over time in minutes for subject 1, y-axis scaled for all parameters noticed on the right. Error rates and blink rate were normed by their maximum.

So blinks might generally only appear directly after stimulus perception. For that, it could in this subject even be possible to (unreliably) predict correct stimulus response solely from blink recordings. Therefore the parameter would be highly correlated with attention to the task.

Many false negative responses (26 %) and no false positive responses to catch trials appeared throughout the whole recording time.

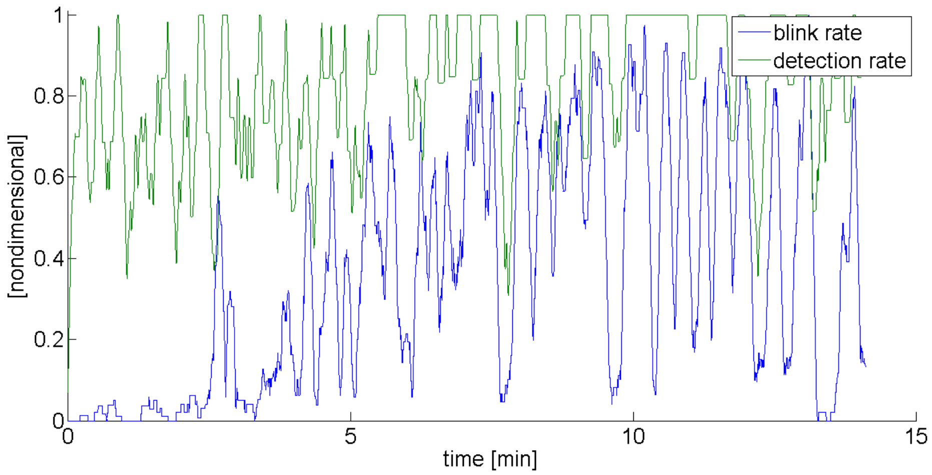

Figure 4.31: Normed detection rate (nondimensional, green) and blink rate (nondimensional, blue) over time in minutes for subject 1, y-axis scaled for parameters noticed on the right. Detection rate and blink rate were normed by their maximum.

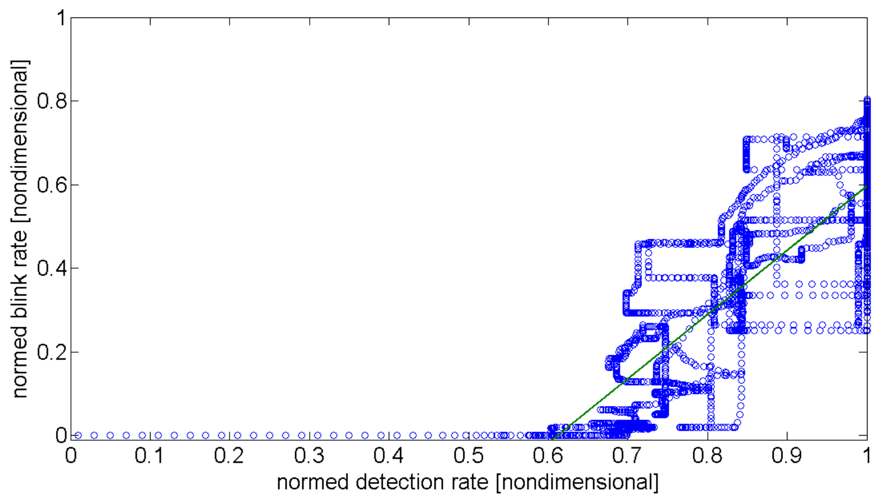

Figure 4.32: Scatterplot for correlation of normed blink rate and normed detection rate (blue circles, both nondimensional). Blink rate and detection rate were normed by their maximum. A regression line (green) with the equation $br = 1.5\ dr - 0.94$ ($br = blink\ rate$ and $dr = detection\ rate$) was calculated.

Subject 2

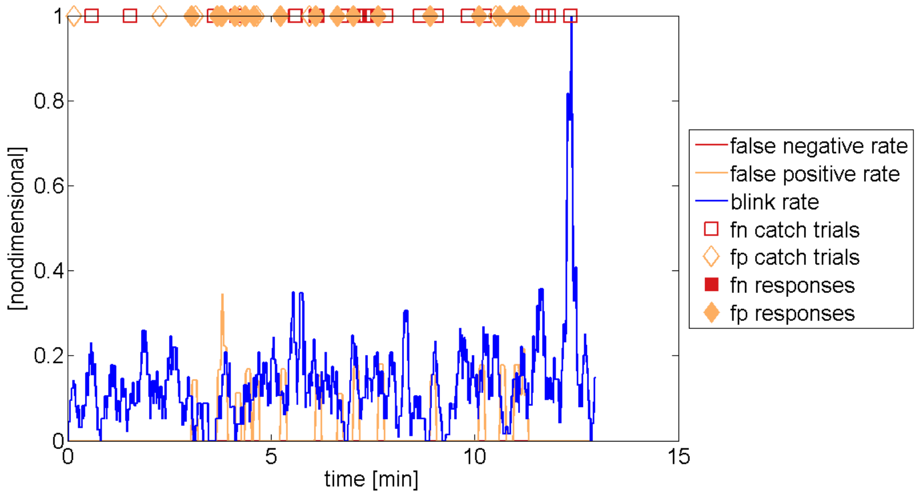

Figure 4.33: Normed error rates (nondimensional, fn: red and fp: orange) and blink rate (nondimensional, blue) over time in minutes for subject 2, y-axis scaled for all parameters noticed on the right. Error rates and blink rate were normed by their maximum.

Figure 4.33 shows error rates and blink rate for subject 2. The calculated correlation coefficient was $R = -0.0745$.
An inacceptable high percentage of false positive responses to catch trials appeared throughout the whole recording session (55 % before and 69 % after the occureence of fatigue waves). No false negative responses to catch trials were given throughout the whole recording session.

Subject 3

Figure 4.34 shows error rates and blink rate for subject 3. The calculated correlation coefficient was $R = 0.0105$.
False positive and false negative respones to catch trials appeared throughout the whole perimetric session, but were both on an acceptable level with 7 % false negative and 19 % false positive responses.

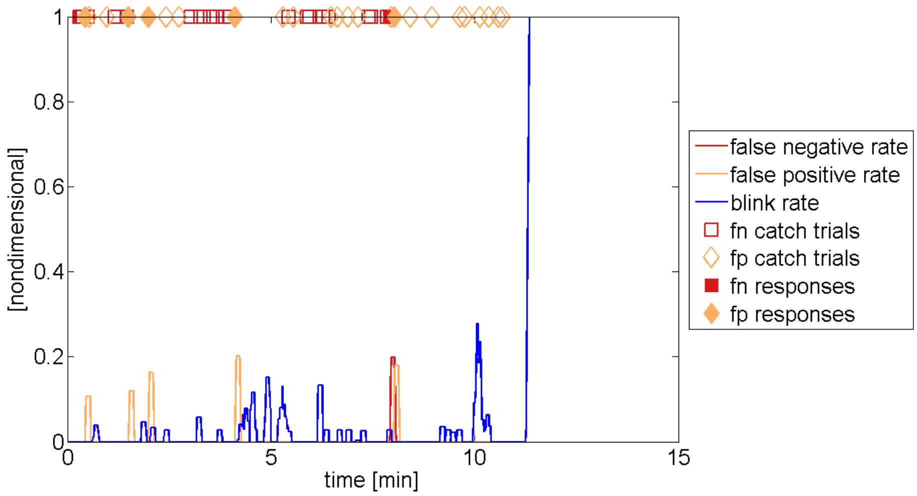

Figure 4.34: Normed error rates (nondimensional, fn: red and fp: orange) and blink rate (nondimensional, blue) over time in minutes for subject 3, y-axis scaled for all parameters noticed on the right. Error rates and blink rate were normed by their maximum.

Subject 4

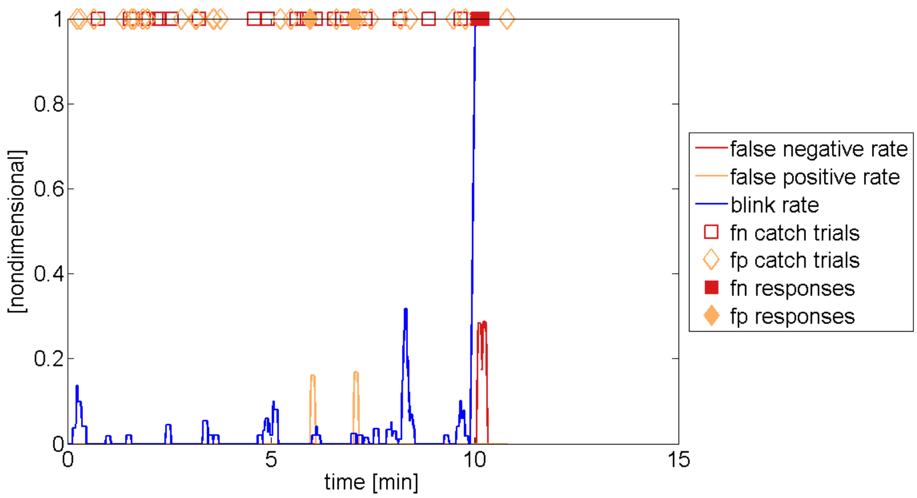

Figure 4.35: Normed error rates (nondimensional, fn: red and fp: orange) and blink rate (nondimensional, blue) over time in minutes for subject 4, y-axis scaled for all parameters noticed on the right. Error rates and blink rate were normed by their maximum.

Figure 4.35 shows error rates and blink rate for subject 4. The calculated correlation coefficient was $R = 0.0994$.

From the beginning of the recording until the appearance of fatigue waves, no false negative and 9 % false positive responses to catch trials were given, which was acceptable. From the appearance of fatigue waves after about eight minutes, no false positive responses to catch trials were given, what was still acceptable. The false negative rate increased to 50 % and for that to an inacceptable level.

Subject 5

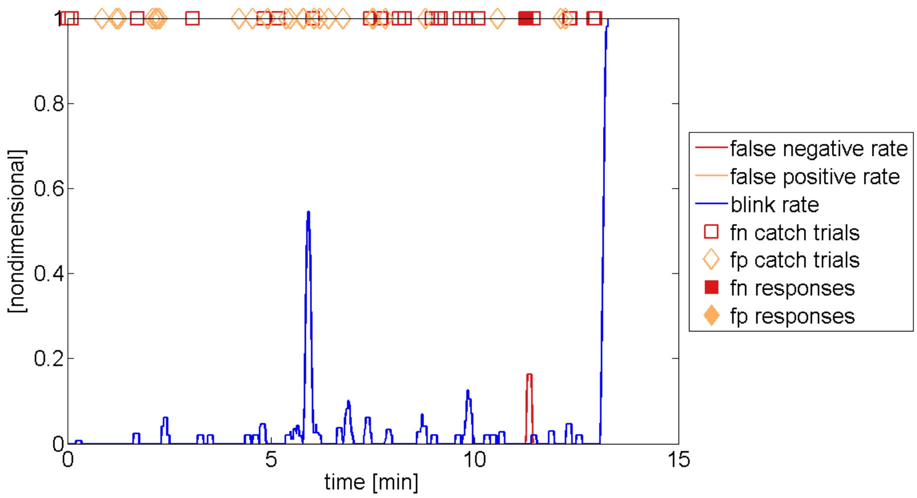

Figure 4.36: Normed error rates (nondimensional, fn: red and fp: orange) and blink rate (nondimensional, blue) over time in minutes for subject 5, y-axis scaled for all parameters noticed on the right. Error rates and blink rate were normed by their maximum.

Figure 4.36 shows error rates and blink rate for subject 5. The calculated correlation coefficient was $R = -0.0143$.

No false positive responses to catch trials and only one false negative response (4 %) to catch trials after the occurence of fatigue appeared throughout the whole perimetric session, which was obviously rated acceptable.

Subject 6

Figure 4.37 shows error rates and blink rate for subject 6. The calculated correlation coefficient was $R = -0.0566$.

False negative responses to catch trials (4 %) appeared as well as false positive responses

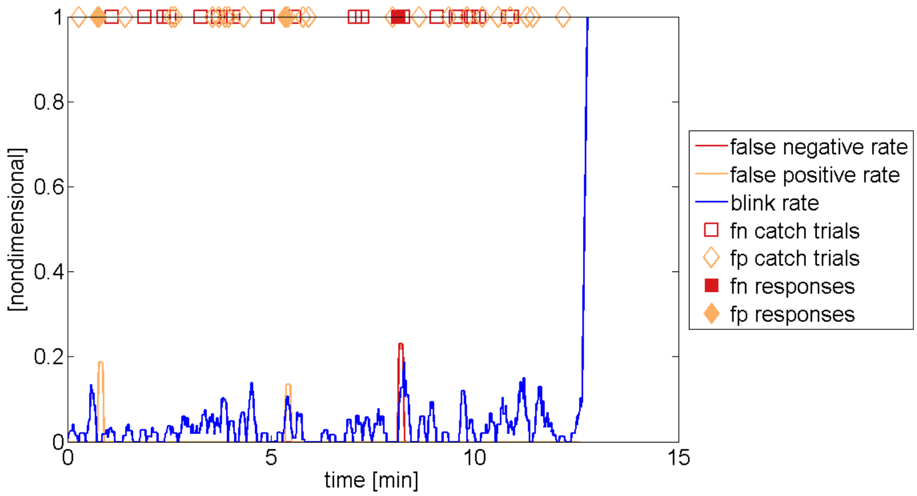

Figure 4.37: Normed error rates (nondimensional, fn: red and fp: orange) and blink rate (nondimensional, blue) over time in minutes for subject 6, y-axis scaled for all parameters noticed on the right. Error rates and blink rate were normed by their maximum.

to catch trials (7 %) throughout the whole recording session. Both false positive and false negative rate were rated as acceptable.

Summary

The blink and error rates have not been discussed for each subject separately, as the same problem appeared for subjects 2, 3, 4, 5 and 6: Blink rates were very low (0 to 3 blinks per minute) and for that, no conclusions in relation to the error rates could be drawn by a description of the individual blink rates. The results for correlation of blink rate and error rates seemed to be not dependable. Even with correct blink detection, a high positive correlation as expected did not appear. That could be due to the relatively low absolute number of catch trials compared to the number of data points registered for pupillography by the IR camera.

A dependable relation between false negative as well as false positive responses to catch trials and the blink rate could not be discovered within this study. That was due to the fact, that a very low number of blinks was detected for subjects 2, 3, 4, 5 and 6 and that the number of false positive and false negative catch trials was also comparatively low.

Significance levels for all correlations have not been calculated, as they are not relevant.

Only significance on an intra-individual basis would have been shown, not for the combined data of all subjects. Due to the small sample size and large inter-individual differences, only a descriptive statistical consideration was performed at this juncture.

Furthermore, subjects 4, 5 and 6 showed a distinct increase of the blink rate at the end of the perimetric session. This could be due to means of physiology. Subjects could, for instance, have suffered from dry eyes towards the end of the recording sessions as rather few blinks were made. Parameters related to the precorneal film and parameters related to attention do not necessarily appear together. Thus effects resulting from parameters related to the precorneal film and to attention could superimpose. For that, effects due to the precorneal film could appear towards the end of the recording session.
Otherwise, the increase appearing for subjects 4, 5 and 6 is rather abrupt and could for that be assumed as caclulation error. Blink rate is by now a very interesting parameter for monitoring attention during perimetry, although we cannot come up with highly reliable results within this study.

4.3 Resume

This study aimed at the identification of parameters for vigilance, attention and cognitive workload. According to data collected from subjects who had to fulfill a visual exploration task, the performance of parameters examined is decribed below.

Fatigue waves, particularly the detail level d10 out of wavelet transformed pupillary data, could work, but have to be defined more precisely. Other parameters showed trends for the combined data of a group of subjects. Those parameters were pupil diameter variability, fixation duration, saccade length and blink duration. The average values (a10) out of wavelet transformed pupillary data for cognitive workload showed additionally significant changes.
Blink rate and the Index of Cognitive Activity (ICA) did, according to the findings of this study, not perform well as indicators for fatigue or cognitive workload.
Horizontal difference between left and right eye and therefore vergence accuracy showed a trend but was hardly usable as an indicator for fatigue, as the effect size was rather small and could therefore be masked by measurement noise.

In special cases of examination, that are represented by perimetry throughout this work, only monocular parameters could be used, as there is only one camera included in the perimeter.

In addition, subjects had to fixate a target. Hence, parameters due to eye movements could not be examined as well. Instead, another parameter could be used, as responses to catch trials had to be given during the perimetric examination: the error rate, that was calculated as the ratio of false positive and false negative responses to catch trials. False negative and false positive responses to catch trials could hardly be related to fatigue waves (d10) and blink rate. No dependable result could be shown as the number of subjects as well as the number of catch trials were comparatively low.

5 Discussion

The results found throughout this study are critically compared to already existing studies in order to show the similarities and differences. Also, possible achievements of this work are discussed. As well, the limitations of this research are mentioned. In addition, an outlook is presented to show what this work and findings could possibly be applicable to and a conclusion is drawn.

5.1 Discussion of Results

In order to find indicators for vigilance, a data set concerning the visual exploration of *Ilya Repin's "Unexpected Visitors"* was recorded and evaluated. A comparison of the results of this experiment to former studies has to be done with caution. None of the former studies has recorded visual exploration of this specific painting with regard to vigilance, attention or cognitive workload. Former studies evaluated data sets for the visual exploration of *Ilya Repin's "Unexpected Visitors"* and for that it can be seen as a standardized painting, Yet, these studies did evaluations with respect to scanpath comparison for different subjects or tasks (Yarbus 1967), (Borji and Itti 2014).

As this painting offered specific characteristics, a comparison to former studies was hardly possible and has to be handled with care. Some parameters for vigilance and attention have not been evaluated for an image viewing task at all.

The blink rate has been evaluated for instance during driving in a simulator (Schleicher et al. 2008) or for subjects sitting silently in front of a blank, neutral wall by EOG (Barbato et al. 2000).

Cognitive workload has been examined for other tasks, such as arithmetical, logical tasks, but not for image viewing (Beatty and Lucero-Wagoner 2000).

Fixation duration was examined while subjects had to drive a monotonous simulator course (Schleicher et al. 2008).

Saccade length depends on the given stimulus or scene, anyway. Hence, no standard values for this parameter can be given and no expectations can be stated.

For this, our results were scarcely comparable to results of former studies and might differ.

For some of the examined subjects, during the visual exloration of *Ilya Repin's "Unexpected Visitors"*, differences between the diameter of left and right eye of up to two pixels were recorded. Resolution errors of one pixel were expected to occur on the edges of the pupil, and these can be added. So measurement errors of up to two pixels in total were expected. For that, the differences between pupil diameter of left and right eye could be due to the inaccuracy of the eye tracker. In addition, the scale of the y-axes varied. This was because the distance from eye tracker to the eyes of the subjects could not be exactly identical in each measure taken. Distance differences of up to 5 cm appeared, as the eye tracker had to be adjusted individually. Because the size of an object in an image depends on the distance the image is recorded from, pupils that would have the same diameter in realitiy can have different diameters in the images taken by the eye tracking camera.
In addition, people might have different sized pupils. Physiological anisocoria of up to 1 mm is considered normal and non-pathological. (Hopf and Koempf 2006). However for all subjects pupillary data showed similar local trends, indicating a good detection accuracy.

The results of this study showed that some parameters examined work well for median values of a group of subjects. Those parameters were pupil diameter variability, the average and detail values out of wavelet transformed pupillary data, fixation duration and saccade length. Only the average values of wavelet transformed pupillary data showed significant changes, for the other parameters mentioned trends appeared.
Blink duration and blink rate, which is already used in some devices monitoring subjects fatigue while driving a car (Schleicher et al. 2008) as well as the Index of Cognitive Activity (ICA) (Marshall 2000) were hardly usable as indicators according to this study, as these parameters did not show any relevant change over time.

Most of the effects could only be observed for the combined data of all subjects but not for the individual. So these parameters worked, but not without fail. In the following reasons for the mismatch between results and expectations as well as the achievements of this study are shown in the context of former studies.

As pupillary oscillations get larger with fatigue (Grünberger et al. 1994), detail values of fatigue waves and the pupil diameter variability were supposed to show stronger oscillation towards the end of the recording time, what they actually did. But before being able to take wavelet transformed pupil diameter data as such an indicator, closer definitions for fatigue waves have to be made and standard values have to be established.

To get dependable results, it would be necessary to know exactly, which coefficient to choose within the wavelet transformation as this depends on the frame rate of the used eye tracking device (Mallat 1987). This would only be possible if a ground truth for eye tracking recordings existed and if this was measured with an eye tracker similar to the one taken for this experiments. Furthermore, there is no detail length for fatigue waves and no minimal change of amplitude defined. Using fatigue waves as an indicator would assume a prior study that connected fatigue waves to an objective fatigue measure.
Waves of pupillary oscillations of about 0.5 to 1 and 4 to 40 seconds duration could be found, also matching Lowenstein's fatigue waves definition (Lowenstein et al. 1963). According to Wilhelm 1999, large changes in pupillary oscillations are called fatigue waves (Wilhelm et al. 1999). Yet a minimum length of oscillation duration is not defined so far.

It is known that changes in pupil dilation accompany effortful cognitive processing (Kahneman 1973) and that pupil size over time can indicate cognitive workload and cognitive activity (Beatty and Lucero-Wagoner 2000). This leads to a parameter due to cognitive activity: The Index of Cognitive Activity (ICA). Wavelet transformation also has an impact on this parameter. A wavelet with specific shape and levels for average and detail values is implemented in the calculation of the ICA. Level and shape are actually due to the frame rate. The originally implemented coefficients erre optimized for another frame rate than ours (Marshall 2000). Up to now, there is no closer definition which coefficients to choose for specific frame rates.

Blink rate was assumed to increase with increasing fatigue (Barbato et al. 2000) (De Padova et al. 2009) and so was blink duration (Tanaka 1999), but both did not perform that way in this study.
A decrease in blink rate occurs while operating on a workstation (Ziemssen et al. 2005). This finding could possibly be assigned to the visual exploration of a painting as it is presented on a monitor. On the other hand, we presented a painting for a duration long enough to make subjects inattentive and tired. These trends compete against one another. This work showed, that blink rate increased for some but not obligatory all subjects over time with supposed increasing fatigue.
In addition, there is a generally high variability in blink rates of several persons as the blink rate is due to tear stability and thus to the precorneal film. The precorneal film is influenced by a variety of circumstances such as thermal factors (low relative humidity, high room temperature) and individual characteristics (e.g. thyroid gland dysfunctions, use of contact

lenses) (Wolkoff et al. 2005). Altogether, the blink rate can, according to this study, neither be used as parameter for fatigue for a group of subjects nor on an individual basis. That could be due to competing effects of fatigue and exploration on a monitor. Blink rate is already used for monitoring fatigue while driving and is seen as a useful parameter for the assessment of fatigue in this situation (Schleicher et al. 2008).

An increase in fixation duration was expected. Cognitive fixations last about 150 - 900 ms and with the incidence of fatigue, basically a higher percentage of overlong and express-fixations appear (Schleicher et al. 2008). Express-fixations were not expected during an experiment of calm visual exploration of a painting, whereas fixation durations of over 900 ms ("staring") were assumed to be an indicator for inattention. And actually there was, according to the results of this study, an increase in fixation duration over all subjects by trend. Nevertheless, this parameter has to be tested further in order to visualize a direct and more detailed relation of fatigue and overlong fixations. In addition, staring could not only occur, when subjects get tired as a matter of principle, but when they get inattentive as a consequence of boredom. That has not to be compulsory towards the end of the recording time, but can occur as soon as subjects get bored by their task, what could already appear in the beginning of the recording session. In addition, subjects could possibly define "their own task" after some time, and for that explore details of the painting attentively, even towards the end of a recording session.

As a binocular eye tracking device was used within this study, we tested the theory that binocular gaze accuracy could decrease with increasing fatigue. But horizontal difference between left and right eye, what shows vergence accuracy, was hardly usable as indicator for fatigue. Reasons for this situation had to be established.

A vergence signal that is the disconjugate eye movement can easiliy be calculated as the Euclidean distance between the position of the left eye and position of the right eye (Jainta and Kapoula 2011). This has been done in former studies, not in consideration of vigilance, but for the determination of the accuracy of binocular vergence. Only the x-coordinate distance was used, as it is crucial for matters of vergence (Jainta and Kapoula 2011).

Vergence accuracy has not been evaluated as an indicator for vigilance formerly but for the first time within this study. We also calculated the vergence accuracy using only the x-coordinate as y-coordinate is not related to vergence (Jainta and Kapoula 2011).

Vergence accuracy was assumed to be a very well-performing parameter for fatigue in the situation of visually exploring a picture, as it is independent from changes in light intensity in contrast to pupillary oscillations (Campbell and Whiteside 1950). In addition, vergence

movements are disrupted by fatigue (Joos et al. 2003). But on the other hand, vergence errors of up to $\pm$ 2°, without diplopia, are common in subjects with normal binocular single vision and errors of 5° are rare but present (Cornell et al. 2003). This circumstance and, in addition, measurement noise seem to have made the effect too small to be relevant.
Concerning vergence accuracy, for future studies it may be interesting to just use distances at the beginning and end of saccades to calculate a metric. In addition, an eye tracker with higher time resolution would be needed for that. Furthermore, the impact of gaze position could be estimated if data of the calibration was available. In this case, a polynom fit could be applied to get rid of correlations between gaze position and vergence accuracy.

The results found during the evaluation of the data set concerning the visual exploration of *Ilya Repin's "Unexpected Visitors"* have been transferred to a second data set concerning perimetry. This data set had already been recorded for another study and was re-evaluated within this work.
Monitoring vigilance during perimetry is possible by pupillography. This is an important development with regard to perimetry as errors in visual field testing occur with patients' increasing fatigue. When errors occur in the case of not perceiving a stimulus, that should have been seen for reasons of fatigue or sleepiness, the luminance difference threshold decreases. Therefore the visual field appears less extended than it realistically is. To get valid results, visual field examinations have to be terminated in time. For that, the quality of visual field testing can be enhanced by terminating a perimetric session as soon as indicators for fatigue occur (Schiefer et al. 2006).
Pupillography has been applied to different tasks in many respects in the recent past. The use of pupillography for monitoring vigilance, fatigue or sleepiness was established with the PST (Wilhelm et al. 1998) licensed by AMTech in 1997. Subjects were placed in front of a camera under scotopic conditions, and had to gaze at a red stimulus emitted by the camera used. So this test cannot be used for any task that takes place under photopic conditions.
Later on, vigilance was monitored during a campimetric task presented on a monitor (Henson et al. 2010). It was shown, that this procedure even works under photopic conditions.
This technique was transferred to a perimetric task, where the camera implemented in a perimeter recorded the subjects' pupil, also under low photopic conditions (Müller 2013).
In the cases described above, subjects had to gaze at a fixations mark. In the first experimental condition of this study, we recorded subjects' pupils under photopic conditions while viewing a painting. Hence, there is no fixation mark, as subjects have to explore the whole painting. Suitable parameters found for these conditions were then transferred to perimetric

data in order to figure out, if they still worked under low photopic conditions with subjects' fixations kept stable. Other parameters due to eye movements such as saccade length and fixation duration could therefore not be transferred. Also binocular parameters such as vergence accuracy could not be transferred to perimetric data, as there is only one camera included in the perimeter, and perimetry is usually performed under monocular conditions.

The data set for the perimetric task showed very low blink rates for five out of six subjects. Some subjects did not blink for several minutes. Hence, an insufficient pupil detection algorithm could be presumed, but according to Ziemssen, blink values like that can actually appear, for instance when operating a workstation (Ziemssen et al. 2005). In this study, subjects had to gaze at a fixation target, so the effect of a very low blink rate could also appear. Corneal clarity decreases in subjects with very low blink rates, what makes pupil detection within an eye tracking examination more difficult. Five out of six subjects showing that low blink rates can be seen as an astounding accumulation. However, data has to be regarded as valid, but in future studies, attention should be paid to this phenomenon.

Fatigue waves were found as expected and in addition it was tried to connect pupil diameter variability, fatigue waves, blink rate and error rate.
Specifically the relation between fatigue waves and error rate has to be handeled with care, as catch trials were limited to 5 % false negative and 5 % false positive. For a total number of 540 questions asked with MoCS strategy, there were only 27 false positive and false negative catch trials, respectively. For comparability reasons that would be 0.075 Hz in average, roughly estimated (average recording time was about 12 minutes). Thus, the temporal resolution for the error rate was, compared to the other parameters (with a frame rate of 20 Hz within the perimetric data), quite rough. Therefore, a correlation analysis might not lead to relevant information. According to the small number of catch trials, false positive and false negative rate were combined as error rate for correlation analysis. Alternatively, with more catch trials implemented, a discrimination between false positives and false negatives could be possible, what could lead to more valuable results.
Correlations were calculated for the error rate which was made a (pseudo-)steady parameter by an enlargement of the data points and sliding window technique. That did not lead to valuable results, most probably due to considerable differences between the frame rate of pupil recordings and catch trials (see above). In addition, all catch trials and responses to catch trials were displayed and described. This method seemed to be more promising, although an increase in false responses to catch trials with increasing fatigue was expected

but did not appear for most of the subjects. To get more dependable results, more catch trials should be implemented as mentioned. Obviously, a frequency comparable to the frame rate of 20 Hz is not feasible for catch trials. Nevertheless, about 30 % false positive and false negative catch trials should be implemented in order to achieve a higher absolute number of catch trials. With this percentage, 162 false positive and 162 false negative catch trials would be implemented. This number of catch trials could lead to a selective error rate for false positive and false negative catch trials as well as to dependable results of correlation analyses. However, connection of fatigue waves, blink rate and error rate still seems to be very interesting and should be investigated in detail within upcoming studies.

5.2 Limitations of Study Design and Methodology

This research project is limited for different reasons:

First of all, the sample size of 20 subjects (valid data: 19/20) for visual exploration of *Ilya Repin's "Unexpected Visitors"* and particularly 8 subjects (valid data: 6/8) for the re-analysis of perimetric data was quite small. For that, this work has to be rated as a pilot study. More subjects could lead to a more distinct identification of effects.

Furthermore, only one kind of eye tracking device was used: The Eye Tribe tracking system for the visual exploration task. Perimetric data was recorded with the camera integrated in the OCTOPUS 900 perimeter. Different types of eye tracking devices that work in different distances from eyes and object of interest should be used to validate the findings within this study. The spatial resolution of eye tracking devices is restricted, so measurement noise and therefore inaccuracy of the measures cannot completely be eliminated.
In addition, a 30 Hz recording rate (for The Eye Tribe tracking system) is quite slow considering the movement speed of the eye.
The Eye Tribe tracking system was used to make sure robust measurement values. Other eye tracking systems would additionally present a picture of the tracked eye and pupil what could have been used to calculate the pupil diameter in a metric unit instead of pixels and also would have allowed for quality control of the image quality. This could only be done with the Eye Tribe tracking system if a defined calibration pattern would have been measured, but that has not been done. An average pupil diameter could not be stated as the pupil diameter is age-related and shows a high inter-individual variability (Loewenfeld 1979).

When applying the parameters to real-time monitoring of vigilance, normal values for the individual necessarily would have to be assessed first.

The subjects had to fulfill just one task while collecting data: The visual examination of *Ilya Repin's "Unexpected Visitors"*. Other tasks should be prepared for more reliable insights. In addition, subjects phoria could have been examined in order to gain more insights. Information about the subjects' phoria could have led to more detailed results concerning the horizontal difference between left and right eye and for that, vergence accuracy as phoria appears more pronounced with increasing fatigue.

The only parameter examined, that has been formerly validated against EEG, is the pupillary oscillation and, for that, pupil size (Berka et al. 2007). The other parameters examined are not validated against EEG, what can be seen as a ground truth, and can therefore be taken as evidence.

It is to mention, that for some parameters such as the vergence accuracy, binocular eye tracking is vital and this parameter cannot be extracted out of monocular data. So monitoring vigilance during perimetric sessions is limited technically, as there is only one included camera available. So for this special case of examination, only monocular parameters can be used.

Combined methods such as MSLT, MWT, CFF-Test, ECP, TAP and SSS (presented in chapter *2.4.2 Monitoring Vigilance*) are not, but could additionally be performed to create a ground truth, although each of them is not without failure. For instance the MSLT, MWT and SSS show a lack in objectiveness, for other methods, such as the ECP, a consistent test was never developed (Weeß et al. 2000).

As is known, there are effects of visual and verbal presentation on cognitive load in vigilance: The magnitudes of pupil response are greater for auditory presentation than for visual presentation (Klingner et al. 2011). During the visual exploration of *Ilya Repin's "Unexpected Visitors"*, the instructions have been given aurally, which may have affected the workload peak at the beginning of those measures. To keep that affection small, visual instructions could have been given.

In addition, there are some physiological characteristics that have to be taken into account: First, there are daytime variations in central nervous system activation and they have an

impact on pupil size and pupillary oscillations (Wilhelm et al. 2001). So, more reliable results could have been generated, if all subjects were examined during the same times of day, what was not possible.
Furthermore, errors in binocular fixation are common, what could affect the examination of binocular vergence. Vergence errors of up to $\pm 2^\circ$, without diplopia, are common in subjects with normal binocular single vision. Errors of 5° are rare but present (Cornell et al. 2003). This could affenct the examination of vergence accuracy, respectively show that the occuring effect is too small to separate it from common fixation errors.
Finally, spontaneous pupillary oscillations in light relate to light intensity, as light-induced pupillary oscillations correlate positively with light intensity (Warga et al. 2009). Light intensity has not been measured throughout this study. Although, it is to assume, that there are no affections to the results as the illumination levels were constant for each subject.

5.3 Conclusion

Altogether, it is to state, that the aim of finding indicators for vigilance, attention and cognitive workload can be seen as reached with this work. Parameters that could work, although not without fail, have been identified with pupil diameter variability, saccade length, fixation duration and fatigue waves. Only the latter parameter has the potential to be applied to perimetric conditions.

5.4 Outlook

Overall the distribution of attention - just as the spatial distribution of gaze - seems to show large interindividual variation and probably also large variance between repetitions. This makes them hard to use for assessing fitness to perform a certain task on an individual level. The assessed average effects could be used e.g. to weight heatmaps of gaze behavior. Areas with a lot of attention (not necessarily with a lot of gaze) would be weighted higher. Since heatmaps are usually averaged over about 30 subjects, the average effect should become clear. For instance attention during viewing a "Wimmelbild" could be assessed by heatmaps: Probably a lot of attention to the searched and interesting objects and only few attention elsewhere could be seen.

Monitoring vigilance during perimetric sessions with additional implemented catch trials and MoCS strategy is highly promising. Fatigue waves have to be defined more precisely and could be connected to the error rate. With more false positive and false negative catch trials implemented, for instance, 30 % each, the error rate could be calculated as an objective measure for fatigue. Fatigue waves could be separated into false positive and false negative rates in order to get more valuable results. Fatigue waves could be determined by a correlation of the error rates (false positive and false negative) and the detail level out of wavelet transformed data. In addition, a correlation of the values for the detail level out of wavelet transformed data and EEG data, which is an objective indicator for fatigue (Berka et al. 2007), could be performed in order to dependably identify fatigue waves. Thus, the quality of visual field testing could be enhanced by terminating a perimetric session as soon as indicators of fatigue occur.

To validate the trends found throughout this study, another study with more subjects should be performed. Here, also different types of eye tracking devices could be used and different tasks could be given to the subjects.
Some things mentioned above in chapter *5.2 Limitations of Study Design and Methodology*, such as daytime variations, the effect of visual and verbal presentation, room illumination and the collection of more data (age, ametropia, phoria etc.) should be considered in such a bigger study.

Bibliography

Addison PS. 2002. The Illustrated Wavelet Transform Handbook. Introductory Theory and Applications in Science, Engineering, Medicine and Finance. Bristol: IOP Publishing Ltd.:2–3.

Barbato G, Ficca G, Muscettola G, Fichele M, Beatrice M, Rinaldi F. 2000. Diurnal variation in spontaneous eye-blink rate. Psychiatry Research, 93:145–151.

Beatty J, Lucero-Wagoner B. 2000. The Pupillary System. In: Cacioppo, JT, Tassinary, LG, Berntson, GG (Eds): Handbook of Psychophysiology 2nd edition New York: Cambridge University Press:142–162.

Bentivoglio AR, Bressman SB, Cassetta E, Carretta D, Tonali P, Albanese A. 1997. Analysis of Blink Rate Patterns in Normal Subjects. Movement Disorders, 12(6):1028–1034.

Berka C, Levendowski DJ, Lumicao MN, Yau A, Davis G, Zivkovic VT, Olmstead RE, Tremoulet PD, Craven PL. 2007. EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory Tasks. Aviation, Space, and Environmental Medicine, 78(5):231–244.

Bleuler E. 1916 / 1983. Lehrbuch der Psychiatrie. 15th edition. Berlin: Springer:76–77.

Borji A, Itti L. 2014. Defending Yarbus : Eye movements reveal observers' task. Journal of Vision, 14(3):29:1–22.

Campbell FW, Whiteside TCD. 1950. Induced Pupillary Oscillations. British Journal of Ophthalmology, 34(3):180–189.

Canisius S, Penzel T. 2007. Vigilance Monitoring - Review and practical aspects. Biomedizinische Technik, 52:77–82.

Carrasco M. 2011. Visual attention: the past 25 years. Vision Research, 51(13):1484–1525.

Chaudhuri A, Behan P. 2004. Fatigue in neurological disorders. Lancet, 363:978–988.

Cornell ED, Macdougall HG, Predebon J, Curthoys IS. 2003. Errors of binocular fixation are common in normal subjects during natural conditions. Optometry and Vision Science, 80(11):764–771.

De Padova V, Barbato G, Conte F, Ficca G. 2009. Diurnal variation of spontaneous eye blink rate in the elderly and its relationships with sleepiness and arousal. Neurosci Lett, 463(1):40–43.

Delabarre EB. 1898. A method of recording eye movements. American Journal of Psychology, 9(4):572–574.

Di Stasi LL, McCamy MB, Macknik SL, Mankin J, Hooft N, Catena A, Martinez-Conde S. 2013. Saccadic eye movement metrics reflect surgical residents' fatigue. Annals of Surgery, 259(4):824–833.

Diamond MR, Ross J, Morrone MC. 2000. Extraretinal Control of Saccadic Suppression. The Journal of Neuroscience, 20(9):3449–3455.

Dodge R, Cline TS. 1901. The angle velocity of eye movements. Psychological Review, 8(2):145–157.

Duchowski AT. 2007. Eye Tracking Methodology. Theory and Practice. 2nd edition. London: Springer.

Endres B. 2009. Einfluss neurologischer Krankheitsbilder auf die Vigilanz gemessen mit dem Pupillographischen Schläfrigkeitstest. PhD thesis, Johannes Gutenberg University Mainz.

Grünberger J, Linzmayer L, Grünberger M, Saletu B. 1994. Eine neue Methode zur Messung der zentralen Aktivierung: Fourieranalyse der Pupillenoszillationen bei depressiven Patienten. Klinische Wochenschrift, 106(6):164–170.

Head H. 1923. The conception of nervous and mental energy. II. Vigilance: A physiological state of the nervous system. British Journal of Psychology, 14:125–147.

Henson DB, Emuh T, Royal M, Hospital E, Road O. 2010. Monitoring Vigilance During Perimetry with Pupillography. IOVS, 44:1–11.

Hess EH, Polt JM. 1964. Pupil Size in Relation to Mental Activity during Simple Problem-Solving. Science, 143:1190–1192.

Holmqvist K, Nystroem M, Andersson R, Dewhurst R, Jarodzka H, van de Weijer J. 2011. Eye Tracking. A Comprehensive Guide to Methods and Measures. New York: Oxford University Press:9–11.

Hopf HC, Koempf D. 2006. Erkrankungen der Hirnnerven. Stuttgart: Thieme:55.

Huey EB. 1898. Preliminary experiments in the physiology and psychology of reading. American Journal of Psychology, 9:575–586.

Jainta S, Kapoula Z. 2011. Dyslexic children are confronted with unstable binocular fixation while reading. PloS one, 6(4):e18694.

James W. 1890. The Principles of Psychology. New York: Henry Holt and Company:403–404.

Joos M, Rötting M, Velichkovsky BM. 2003. Bewegungen des menschlichen Auges: Fakten, Methoden und innovative Anwendungen. Psycholinguistik: Ein internationales Handbuch Berlin and New York: De Gruyter:142–168.

Kahneman D. 1973. Attention and Effort. Englewood Cliffs, NJ: Prentice Hall:17–24.

Kaufmann HE. 2004a. Strabismus. 3rd edition. Stuttgart: Thieme:158–159.

Kaufmann HE. 2004b. Strabismus. 3rd edition. Stuttgart: Thieme:168–169.

Klingner J, Tversky B, Hanrahan P. 2011. Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. Psychophysiology, 48(3):323–332.

Korczyn A. 1987. The pupil and vigilance. Functional Neurology, 2(4):539–544.

Kraemer S, Danker-Hopfe H, Dorn H, Schmidt A, Ehlert I, Herrmann WM. 2000. Time-of-Day Variations of Indicators of Attention: Assessment of Sleepiness. Biological Psychiatry, 48:1069–1080.

Lang A, Gaertner C, Ghassemi E, Yang Q, Orssaud C, Kapoula Z. 2014. Saccade-vergence properties remain more stable over short-time repetition under overlap than under gap task: a preliminary study. Frontiers in Human Neuroscience, 8:372:1–9.

Loewenfeld IE. 1979. Pupillary changes related to age. In: Thompson H S et al. (Eds.): Topics in Neuro-Ophthalmology. Baltimore: Williams and Wilkins:124–150.

Lowenstein O, Feinberg R, Loewenfeld IE. 1963. Pupillary movements during acute and chronic fatigue. A new test for the objective evaluation of tiredness. IOVS, 2(2):138–157.

Mallat SG. 1987. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. GRASP LAB: General Robotics and Active Sensory Perception Laboratory Technical report, University of Pennsylvania, Dept of Computer and Information Science:1–28.

Marshall SP. 2000. Method and Apparatus for Eye Tracking and Monitoring Pupil Dilation to Evaluate Cognitive Activity. U.S. Patent 6,090,051.

Marshall SP. 2002. The Index of Cognitive Activity: Measuring Cognitive Workload. IEEE 7th Human Factors Meeting Scottsdale Arizona:5–9.

Moses RE. 1981. Adler's Physiology of the eye. Clinical application. Mosby:1–15.

Müller M. 2013. Vigilanzüberwachung während automatisierter, statischer Perimetrie. Master thesis, University of Tuebingen:22–26.

Müller M, Ungewiss J, Kasneci E, Rosenstiel W, Schiefer U. 2014. Monitoring vigilance during automated static perimetry. Poster presentation. 21st International Visual Field & Imaging Symposium, New York, September 9 - 12 2014.

Rantanen EM, Goldberg JH. 1999. The effect of mental workload on the visual field size and shape. Ergonomics, 42(6):816–834.

Russo M, Thomas M, Thorne D, Sing H, Redmond D, Rowland L, Johnson D, Hall S, Krichmar J, Balkin T. 2003. Oculomotor impairment during chronic partial sleep deprivation. Clinical Neurophysiology, 114:723–736.

Schiefer U, Pätzold J, Wabbels B, Dannheim F. 2006. Konventionelle Perimetrie. Teil 4: Statische Perimetrie: Befundauswertung – Indizes – Verlaufskontrolle – Perimetrie im Kindesalter. Der Ophthalmologe, 103(3):235–256.

Schleicher R, Galley N, Briest S, Galley L. 2008. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? Ergonomics, 51(7):982–1010.

Shapiro C, Flanigan M, Fleming J, Morehouse R, Moscovitch A, Plamondon J, Reinish R, Devins G. 2002. Development of an adjective checklist to measure five faces of fatigue and sleepiness. Data from a national survey of insomniacs. Journal of Psychosomatic Research, 52:467–473.

Sweller J. 1988. Cognitive Load During Problem Solving: Effects on Learning. Cognitive Science, 12:257–285.

Tanaka Y. 1999. Arousal level and blink activity. Shinrigaku Kenkyu: The Japanese Journal of Psychology, 70(1):1–8.

The Eye Tribe Aps. 2013. Products - The Eye Tribe: The Tech Specs. Copenhagen. Website. Online: https://theeyetribe.com/products/, latest access: 2015-02-20.

Turpin A, Artes PH, McKendrick AM. 2012. The Open Perimetry Interface : An enabling tool for clinical visual psychophysics. Journal of Vision, 12(11):22:1–5.

Vernet M, Kapoula Z. 2009. Binocular motor coordination during saccades and fixations while reading: A magnitude and time analysis. Journal of Vision, 9:1–13.

Volkmann FC. 1986. Human visual suppression. Vision Research, 26(9):1401–1416.

Warga M, Lüdtke H, Wilhelm H, Wilhelm B. 2009. How do spontaneous pupillary oscillations in light relate to light intensity? Vision Research, 49(3):295–300.

Weeks M. 2011. Digital Signal Processing Using MATLAB & Wavelets. 2nd edition. Boston: Jones and Bartlett Publishers, LLC:271–272.

Weeß HG. 2006. Multifaktorielles Symptom vieler Einschlafstörungen. Tagesschläfrigkeit birgt ein hohes Risiko. Der Klinikarzt, 35:145–152.

Weeß HG, Lund R, Gresele C, Boehning W, Sauter C, Steinberg R. 1998. Vigilanz, Einschlafneigung, Daueraufmerksamkeit, Müdigkeit, Schläfrigkeit. Die Messung müdigkeitsbezogener Prozesse bei Hypersomnien. Theoretische Grundlagen. Somnologie, 2:32–41.

Weeß HG, Sauter C, Geisler P, Boehning W, Wilhelm B, Rotte M, Gresele C, Schneider C, Schulz H, Lund R, Steinberg R. 2000. Vigilanz, Einschlafneigung, Daueraufmerksamkeit, Müdigkeit, Schläfrigkeit. Diagnostische Instrumentarien zur Messung müdigkeits- und schläfrigkeitsbezogener Prozesse und deren Gütekriterien. Somnologie, 1:20–38.

Wickens C, Hollands J. 2000. Engineering Psychology and Human Performance. 3rd edition. Saddle River, NJ: Prentice Hall:459.

Wilhelm B, Giedke H, Lüdtke H, Bittner E, Hofmann A, Wilhelm H. 2001. Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. Journal of Sleep Research, 10(1):1–7.

Wilhelm B, Lüdtke H, Wilhelm H. 1999. Spontaneus pupillary oscillations – An objective measure for the level of tonic central nervous activation. Clinical Pharmacology, 18:27–36.

Wilhelm B, Wilhelm H, Lüdtke H, Streicher P, Adler M. 1998. Pupillographic assessment of sleepiness in sleep-deprived healthy subjects. Sleep, 21:258–265.

Wolkoff P, Troiano P, Piccoli B. 2005. Eye complaints in the office environment: precorneal tear film integrity influenced by eye blinking efficiency. Occupational and Environmental Medicine, 62:4–12.

Yang Q, Bucci MP, Kapoula Z. 2002. The latency of saccades, vergence, and combined eye movements in children and in adults. IOVS, 43(9):2939–2949.

Yarbus AL. 1967. Eye Movements and Vision. New York: Plenum Press.

Ziemssen F, Freudenthaler N, Regnery K, Schlote T. 2005. Lidschlagaktivität während der Bildschirmarbeit. Teil 1: Okuläre Beschwerden und pathophysiologische Grundlagen. Der Ophthalmologe, (102):805–811.

A Results of Correlation Analysis for Vergence Accuracy and Gaze Position

Subject 2

Table A.1: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 2. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.1215	0.0227	vac	1.0000	0.0460	0.0315
gp-x	0.1215	1.0000	0.0184	gp-x	0.0460	1.0000	0.1713
gp-y	0.0227	0.0184	1.0000	gp-y	0.0315	0.1713	1.0000

Subject 3

Table A.2: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 3. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.0431	-0.0286	vac	1.0000	-0.0444	-0.0331
gp-x	0.0431	1.0000	0.4720	gp-x	-0.0444	1.0000	0.4017
gp-y	-0.0286	0.4720	1.0000	gp-y	-0.0331	0.4017	1.0000

Subject 4

Table A.3: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 4. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.0129	-0.1429	vac	1.0000	0.0789	-0.0531
gp-x	-0.0129	1.0000	0.2098	gp-x	0.0789	1.0000	0.2282
gp-y	-0.1429	0.2098	1.0000	gp-y	-0.0531	0.2282	1.0000

Subject 6

Table A.4: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 6. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.0546	0.0588	vac	1.0000	-0.0356	0.0662
gp-x	0.0546	1.0000	0.1461	gp-x	-0.0356	1.0000	0.0579
gp-y	0.0588	0.1461	1.0000	gp-y	0.0662	0.0579	1.0000

Subject 8

Table A.5: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 8. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.1431	0.1658	vac	1.0000	-0.2652	0.0986
gp-x	-0.1431	1.0000	-0.0723	gp-x	-0.2652	1.0000	-0.0796
gp-y	0.1658	-0.0723	1.0000	gp-y	0.0986	-0.0796	1.0000

Subject 9

Table A.6: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 9. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.0047	0.0477	vac	1.0000	0.0603	0.1120
gp-x	-0.0047	1.0000	0.1670	gp-x	0.0603	1.0000	0.1568
gp-y	0.0477	0.1670	1.0000	gp-y	0.1120	0.1568	1.0000

Subject 10

Table A.7: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 10. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.0332	-0.1339	vac	1.0000	-0.0924	-0.1164
gp-x	-0.0332	1.0000	-0.0025	gp-x	-0.0924	1.0000	0.0141
gp-y	-0.1339	-0.0025	1.0000	gp-y	-0.1164	0.0141	1.0000

Subject 12

Table A.8: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 12. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.0320	-0.0924	vac	1.0000	-0.0865	-0.0857
gp-x	-0.0320	1.0000	-0.1077	gp-x	-0.0865	1.0000	-0.1394
gp-y	-0.0924	-0.1077	1.0000	gp-y	-0.0857	-0.1394	1.0000

Subject 13

Table A.9: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 13. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.1110	0.0425	vac	1.0000	-0.0044	0.0467
gp-x	0.1110	1.0000	0.1140	gp-x	-0.0044	1.0000	0.0572
gp-y	0.0425	0.1140	1.0000	gp-y	0.0467	0.0572	1.0000

Subject 14

Table A.10: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 14. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.0245	0.0706	vac	1.0000	0.0284	0.0640
gp-x	0.0245	1.0000	-0.0177	gp-x	0.0284	1.0000	-0.1947
gp-y	0.0706	-0.0177	1.0000	gp-y	0.0640	-0.1947	1.0000

Subject 15

Table A.11: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 15. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.0274	0.0625	vac	1.0000	-0.0405	0.0635
gp-x	0.0274	1.0000	0.2263	gp-x	-0.0405	1.0000	0.2021
gp-y	0.0625	0.2263	1.0000	gp-y	0.0635	0.2021	1.0000

Subject 17

Table A.12: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 17. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.1143	-0.0999	vac	1.0000	-0.1664	-0.1211
gp-x	-0.1143	1.0000	0.6185	gp-x	-0.1664	1.0000	0.5143
gp-y	-0.0999	0.6185	1.0000	gp-y	-0.1211	0.5143	1.0000

Subject 18

Table A.13: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 18. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.2090	-0.0751	vac	1.0000	0.1184	-0.0471
gp-x	-0.2090	1.0000	-0.0282	gp-x	0.1184	1.0000	0.0593
gp-y	-0.0751	-0.0282	1.0000	gp-y	-0.0471	0.0593	1.0000

Subject 19

Table A.14: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 19. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.2388	0.0035	vac	1.0000	-0.2392	0.0229
gp-x	0.2388	1.0000	0.1004	gp-x	-0.2392	1.0000	0.1599
gp-y	0.0035	0.1004	1.0000	gp-y	0.0229	0.1599	1.0000

Subject 20

Table A.15: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 20. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.0511	-0.0577	vac	1.0000	-0.1978	-0.1248
gp-x	-0.0511	1.0000	0.1673	gp-x	-0.1978	1.0000	0.3214
gp-y	-0.0577	0.1673	1.0000	gp-y	-0.1248	0.3214	1.0000

Subject 21

Table A.16: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 21. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.0892	-0.2069	vac	1.0000	-0.0737	-0.1872
gp-x	0.0892	1.0000	0.0753	gp-x	-0.0737	1.0000	0.1063
gp-y	-0.2069	0.0753	1.0000	gp-y	-0.1872	0.1063	1.0000

Subject 22

Table A.17: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 22. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.3254	-0.6970	vac	1.0000	0.1166	-0.6402
gp-x	0.3254	1.0000	-0.1196	gp-x	0.1166	1.0000	0.1826
gp-y	-0.6970	-0.1196	1.0000	gp-y	-0.6402	0.1826	1.0000

Subject 23

Table A.18: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 23. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	-0.1443	-0.1672	vac	1.0000	-0.1111	-0.1721
gp-x	-0.1443	1.0000	0.4529	gp-x	-0.1111	1.0000	0.3636
gp-y	-0.1672	0.4529	1.0000	gp-y	-0.1721	0.3636	1.0000

Subject 24

Table A.19: Correlation matrix with Spearman correlation coefficients for vergence accuracy (vac), gaze position for x-coordinate (gp-x) and gaze position for y-coordinate (gp-y) of left and right eye for subject 24. Diagonals are the lines of identity and filled by the values $R = 1.0000$ (highlighted in gray).

	left eye			right eye			
	vac	gp-x	gp-y		vac	gp-x	gp-y
vac	1.0000	0.3427	-0.4538	vac	1.0000	-0.1403	-0.6133
gp-x	0.3427	1.0000	0.0760	gp-x	-0.1403	1.0000	0.4841
gp-y	-0.4538	0.0760	1.0000	gp-y	-0.6133	0.4841	1.0000

Curriculum Vitae

	Personal Details
Name:	Judith Ungewiß
Date of birth:	12.07.1985
Place of birth:	Herrenberg, Germany
	Work Experience
09/2001 - 02/2004	District Office Calw, Germany
	Apprenticeship as Clerk in Public Administration
02/2004 - 08/2004	Social Services Department, District Office Calw, Germany
	Clerk in Public Administration
	Education
09/2004 - 07/2005	Hermann-Gundert-Schule Calw, Germany
	Degree: Fachhochschulreife
09/2005 - 07/2007	Kaufmännische Schule 1 Stuttgart, Germany
	Degree: Abitur
10/2007 - 08/2009	Student of Physics
	Stuttgart University, Germany
09/2009 - 08/2013	Student of Optometry and Audiology
	Aalen University of Applied Sciences, Germany
	Degree: Bachelor of Science
	Topic of Bachelor Thesis: "MATLAB in Augenoptik & Hörakustik"
since 09/2013	Student of Ophthalmic Optics and Psychophysics
	Aalen University of Applied Sciences, Germany
	Degree to be earned: Master of Science
	Awards
11/2013 - 10/2014	Karl-Steinbuch-Grant of the Medien- und Filmgesellschaft Baden-Württemberg for the project "Comparison of Attention Allocation"/ "Methoden zum Vergleich der Aufmerksamkeitszuwendung"