MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Visual exploration of reading in everyday life

Investigating reading within urban spaces with mobile eye-tracking in the real world

verfasst von / submitted by Alina Schnebel, BSc

angestrebter akademischer Grad / in partial fulfillment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021 / Vienna, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Psychologie
Betreut von / Supervisor:	Univ.-Prof. Dr. Helmut Leder
Mitbetreut von / Co-Supervisor:	

Acknowledgements

First of all, my special thanks go to my supervisor, Prof. Dr. Helmut Leder for his support and encouragement, and for making this thesis possible in the first place. Moreover, I would like to thank Kirren Chana for her ideas and feedback as well as her help during the data collection. To both I am very grateful for always taking the time to help me with open questions.

My heartfelt thanks go to my boyfriend, Renaud Heine, for his continuous support through the most difficult stages of this work. Moreover, I would like to thank Una Vardic for proofreading and her helpful comments on an earlier version of this thesis.

Finally, I would like to thank my parents, who supported me in every possible way throughout my studies.

Table of contents

Abstract

Natural viewing behaviour is usually collected in the laboratory where pictures or videos appear for a few seconds on a computer screen. However, what does this say about the gaze behaviour in real world conditions? This thesis investigates how people encounter text in an everyday setting. Research in the laboratory was already able to show that text generally draws attention, by using pictures of real-world scenes (Cerf, Frady & Koch, 2009). The implemented study design extends from a free exploration paradigm at the Danube canal in Vienna containing urban art (Mitschke, Goller & Leder, 2017). In this study a mixed method design that combined mobile eye tracking and questionnaires was used. First, a mobile eye-tracking field study was conducted. In a follow-up questionnaire, participants were asked about the recall, recognition, and the aesthetic evaluations of text elements that they encountered during the walk. The results of the mobile eye tracking analysis showed that text - with 24.37% of the mean viewing time - drew the most attention, followed by people and nature. These results are significant, since participants were not instructed to look at text.

1 .Introduction

When inspecting real world scenes, human observers continually shift their gaze to receive information (Wang & Pomplun, 2012). This visual information is essential for their interaction with the environment. A particular element of this would be text detection, which is crucial in gathering information for example, from street signs. Humans continuously move in complex visual environments that consist of many different objects. This is especially the case in the shopping streets of larger cities, where a higher density of different stimuli can be encountered. Cerf et al. (2009) investigated eye fixation in natural images in the laboratory in which they could show that eyes are rapidly and strongly attracted to faces and texts, compared to the slower and fewer fixations to mobile phones. In terms of attracting gaze, text was almost as effective as faces. Wang and Pomplun (2012) paired text objects with matching control objects and could show that texts drew more attention than other non-text objects and control regions. It should be noted that these studies were conducted in the laboratory replicating natural circumstances.

The basis of this study design expands on a free exploration paradigm in a natural environment containing urban art (Mitschke et al., 2017). The researchers tested 12 participants in a free exploration walk along the Danube canal in Vienna equipped with a mobile eye tracker. They used portable eye-tracking glasses to record what the subjects looked at. The researchers analysed the viewing behaviour towards graffiti and sculptures. They found that measuring viewing time with eye movement parameters quantified aesthetic value on a physiological level. Mitschke et al. (2017) could show that participants spent up to 50% of the overall fixation time exploring aesthetic objects and artworks. In the following thesis, this paradigm will be used to consider the beauty of text and investigate where people look when walking along a commercial shopping street. The participants' eye movements will be tracked in a free-exploration task along Mariahilferstrasse, one of the biggest and most well-known commercial shopping streets in Vienna. The focus of the study will be to answer the question whether text attracts more attention than other objects in a daily situation. Additionally, it will be shown whether there is a predictive effect of aesthetic value on viewing behaviour and memorability of text. Moreover, the implications of the use of mobile eye-tracking glasses in a real-world setting are discussed.

The research on aesthetic experience mainly concentrates on museums as context but ignores aesthetic decisions in everyday life. Research has previously demonstrated that the museum context - artwork presented in a museum as opposed to in a laboratory - has an impact on the liking and viewing time of artworks (Brieber et al., 2015; Brieber et al., 2014). Mitschke et al. (2017) could show that aesthetic liking and interest for certain views predicted gaze behaviour. The results of Mitschke et al. (2017) also replicated the findings in which texts had been shown to capture gaze in natural scenes and laboratory settings (Cerf et al., 2008).

When, what and how long people look at something can give insights into perceptual and cognitive processes. The selective visual attention is an important function, that directs one’s gaze rapidly towards informative objects in the visual environment. Most eye-tracking research has been conducted in laboratory settings and whilst they have delivered remarkable results, it is important to also concentrate on real life environments to gain information about the viewing behaviour in natural, real-world settings.

It should be pointed out, that it is possible for a person to fixate on a region while paying attention to something else. Posner et. al. (1980) showed that participants could allocate attention independent of eye position in simple tasks in which they had to hold their eyes on a fixation target while other stimuli were presented elsewhere in the visual field. In general, researchers assume that the fixation correlates tightly with the covert attention (Einhäuser et. al., 2006; Parkhurst et. al. 2002). The so-called overt attention can be measured with the help of an eye-tracker. Eye movements provide an unobtrusive, sensitive, real-time behavioural index of ongoing visual and cognitive processing (Henderson, 2003).

The worldwide pandemic of the coronavirus disease 2019 universally affected every aspect of life and therefore also this study. The coronavirus disease is a pulmonary viral disease that broke out in the end of 2019 and has affected life globally ever since. The Austrian government imposed a variety of counter measures such as mandatory testing, face masks, personal distance and lockdowns. These measures had to be taken into account during the testing phase.

1.1 Text

Text, as the physical incarnation of language, is one of the basic tools for preserving information (Jaderberg et al., 2015). Cities, compared to most other environments and landscapes, tend to be more permeated with text among other things, on walls, signs, monuments and shops. Little is known about natural reading experiences in daily activities. The way people encounter texts in everyday life is moving away from traditional reading in books and onto the streets, into the real world. Text reading is part of visually exploring and understanding the environment. It is a cognitive process that helps the reader to understand the meaning of words and receive information (Kendeou et al., 2014).

Text in urban spaces is encountered in many different fonts and styles, with different luminosity, colours, sizes, languages and orientations. Additionally, text can appear in the form of street signs, on advertisement billboards, commercial shop signs but also in the form of graffiti, posters and stickers. All these different text elements have different functions. Some have a concrete function by pointing the way, while others have the goal to draw as much attention as possible for commercial purposes. The streets in Vienna, especially the shopping streets, present large amounts and variations of text and graphic elements. In the real world, text rarely appears in isolation, but rather together with other objects.

Besides text, people also encounter other objects in public spaces. For this reason, the sole focus was not only on text, but also on other categories, including people, architecture, nature, advertisement, street signs, graphic elements, traffic, shop signs and other text in form of stickers or graffiti (see Figure 1).

1.2 Empirical aesthetics

Since Fechner (1876) first established the field of experimental aesthetics, a lot of research has been conducted in that field to provide more detailed descriptions and insights into aesthetic experience. His “Vorschule der Ästhetik” (Propaedeutics of Aesthetics) offers the basis for an experimental approach to art and aesthetics. Empirical aesthetics is interested in the interaction between psychology and neuroscience, the goal being to understand how people experience and evaluate objects aesthetically.

Aesthetic research is not only limited to art but also to non-artistic objects. Pearce et. al. (2016) pointed out that art and aesthetics are not identical but overlap conceptually and historically. Therefore, aesthetic research is something that is not only limited to the museum context. Art as one example of aesthetic experience is encountered in different situations and parts of the daily life. Previous research suggests that the context in which participants encounter artworks has an influence on the aesthetic experience of it (Brieber at al., 2014). The researchers compared the aesthetic experience of artworks in the museum and laboratory contexts. They could show that participants liked artworks more when they viewed them in a museum, compared to in a laboratory. Moreover, they suggested that the laboratory context shortens the experience of art. That should be considered critical since most of the research in empirical aesthetics is carried out in the laboratory. A lot of research has been carried out to understand the connection between aesthetics and art. However, art is only one example of aesthetic research and therefore it is important to note that aesthetic experience includes much more than that. Research on aesthetic responses to art gives interesting insights but does not encompass the whole concept of aesthetic experience. In this thesis the focus is not on artistic objects, but on the aesthetical experience while walking in a public space. Therefore, the focus lies on different appearances of text elements and the aesthetic experience while encountering those.

The research topic for this thesis focuses on an everyday life setting and therefore employs a field experiment rather than a laboratory session. A large city such as Vienna is filled with numerous graphical elements. Huerta (2010) suggested that walking in the city is an aesthetic experience, where an encounter with an urban signboard elicits an experience that can be compared to certain perceptions activated while visiting a museum. Commercial streets are designed to be appealing to visitors and signboards play an important factor in this. Signboards can be an attractive element that engages pedestrians' attention and serve as an advertisement medium for stores.

As aforementioned, it is possible for participants to look at one subject but concentrate on another or in reverse to not look at something but still pay attention to it. The attraction of gaze is not necessarily a positive signal – meaning that an object is well liked by its observer, which is why the study also entailed questionnaires with different rating scales to find out more about the aesthetic value of the single signs. Holmes and Zanker (2011) suggested that gaze fixation time can be used as an indicator for aesthetic preferences. Therefore, the areas that participants focused on, the so-called regions of interest, must be quantified.

What determines how people direct their looks and how they engage with text in an everyday setting? To answer this question, studies need to migrate from laboratory sessions to field experiments. For this study, the focus was placed on shop signs, given that these are the most common and largest form of text in a commercial shopping street. However, other types of text were also included, such as graffiti or stickers.

1.3 Eye-tracking in real world settings

Most of what is known about how people look at natural scenes is based on the ‘free viewing’ of static images presented for a few seconds on a screen in a laboratory with restrained head positions. The visual behaviour of the participants provides an objective and measurable indication of cognitive processes. Eye-tracking technology has been used extensively in empirical research and enables researchers to measure the visual behaviour. Eye movements during reading were already recorded in the 19th century through direct observation. Louis Javal (1879) was the first one to observe, that eye movements during reading were not smooth, but rather a combination of fixations (Huey, 1898). Eye-tracking evaluates the movement of the eyes, which consists of fixations (i.e. pause in a particular scene area) and saccades (i.e. movement to the next fixation). Until recently, the use of eye-tracking systems was somewhat limited, because its use was complex, time consuming and expensive (Rosa, 2015). Some studies have already collected data for specific tasks in natural settings e.g. for driving (Land & Tatler, 2001) or food preparation (Land & Hayhow, 2001).

Laboratory experiments aim to maximise control over all variables. Although these studies have delivered important results, they almost exclusively concentrate on a two-dimensional copy of the real world, presented on a computer screen. Such controlled methods are important to isolate and understand visual and cognitive processes, but the results from these methods cannot be generalised on natural scenes. Eye movements are necessary, since viewers only see clearly with high acuity and colour sensitivity in the fovea, which covers around 2 ° in the centre of the vision (Rayner, 1998; 2009).

In the attempt to perceive their environment and absorb information, humans continually shift their gaze and in doing so perform an average of three eye movements per second (Cristino & Baddeley, 2009). Gaze behaviour describes the pattern of eye movements for an individual, which includes both saccades and fixations. It is therefore critical for efficient, timely acquisition of visual information and is an ongoing cognitive task. To understand eye movement control, one must concentrate on where the fixation position tends to centre and for how long this remains on a particular location (Underwood, 1998).

Foulsham et al. (2011) investigated the difference between viewing video clips of the "natural environment" in the laboratory and participants walking around in the natural environment with mobile eye-tracking glasses. They were able to find some notable differences, pedestrians for example, were fixated on more often when viewed in the video compared to during an outside walk. This offers a valid argument why it is crucial to measure eye movement in natural environments and not only with the help of a screen in the laboratory. Apart from that a human's visual angle is larger than what can be projected on a computer screen and eye movement is accompanied by head and body movement, both of which cannot be captured in seated, stationary laboratory experiments. Mobile eye-tracking allows the viewer to move and interact freely with the environment, without restrictions. It provides precise data on the sensory viewing processes.

1.4 Field study

To establish a foundation for testing cognitive processes in laboratories that can predict real life phenomena, Kingstone et al. (2008) pointed out that studies need to be grounded in real world observations which is why a field study was chosen for this thesis. Here, eye-tracking was used to measure quantitative physiological behaviour and was combined with questionnaires and ratings for psychological measurements. Since there are only a few studies that focus on everyday life in an outdoor setting, this thesis also aims to explore possible benefits and challenges of using eye-tracking in a real-life situation away from the laboratory. Therefore, a naturalistic, unconstrained task was conducted with a mobile eye-tracking head set in an everyday life setting.

1.5 Aim and research questions

The goal of this master thesis is to investigate how people engage and explore text in everyday settings. In order to examine the underlying visual processes in natural circumstances, a mixed method approach was adopted. Viewing time and location was tracked with an eye- tracking headset; this was followed by a memory recall and recognition task, an evaluation rating task, and a Language History Questionnaire (Li et al., 2020). Since a field study was chosen, the research approach was exploratory and therefore guided by open research questions. The interest lies in whether there is a predictive effect of aesthetic value on viewing behaviour and memorability of text.

Research questions:

Which objects capture gaze frequently?

How do people engage and explore text in an everyday setting?

Does the aesthetic value of the text have an impact on the viewing behaviour?

Does the aesthetic value of the text have an impact on the memorability of text?

2. Method

A mixed method was chosen to exhaust the potential of the study. Since the data was collected in a real, everyday environment, it is also relevant to note that the researchers did not attempt nor found it possible to control single variables, for example, traffic or the number of people present. The author used quantitative physiological - in form of eye-tracking - and psychological measurements - in form of rating procedure and different questionnaires - to relate gaze patterns to the aesthetic evaluation and memory of text.

2.1 Participants

In total, 39 participants (25 female), with a mean age of 24.3 (SD= 3.5), took part in the experiment. One participant had to be excluded due to incomplete data. However, only the eye tracking data of 24 (16 female) participants, with the mean age of 23.7 (SD=3.2) was analysed. The videos of 15 participants had to be excluded due to technical issues. There was a problem with the portable hardware where the videos were not saved properly and therefore could not be used for the analysis of the annotations. This occurred during the testing process and wasn't noticeable until afterwards.

For this study, only German speaking participants were tested. All participants reported normal or corrected to normal vision. They were naive to the purpose of the study and received a monetary compensation of 10€ after completing the second follow up questionnaire.

2.2 Apparatus

The Pupil Core headset was used for this particular field study. The Pupil labs headset is lightweight and adjustable. It consists of a world camera and two eye cameras and allows a quick calibration. One limitation of real-world eye tracking is sunlight, as exposure to sunlight can lead to inaccurate tracking. As a preventative measure, testing was conducted early in the morning and all participants were instructed to wear baseball caps to provide further shade over the eyes. The output of the mobile eye-tracking is a video of the first world scene camera with an estimated gaze point. Manual annotation was used to assign a code to the objects that were fixated on in the videos.

2.3 Eye-tracking at Mariahilferstrasse

For this study, the researchers chose one of the biggest commercial shopping streets in Vienna, the Mariahilferstrasse. For an overview of the part of the Mariahilferstrasse that was chosen, see Figure 2. In just a short stretch, there is an array of text elements, including street signs, shop signs, advertisements, but also text in the form of stickers and graffiti.

The study was conducted during the Covid 19 pandemic. The restrictions imposed by the government also affected the study and the amount of people that could be tested. Since the study was conducted outdoors it could be carried out in compliance with hygiene and distance regulations. All participants had to bring a 3G certificate – proof that they were either tested, vaccinated or recovered - and had to wear a FFP2 facemask during the calibration. These restrictions caused by the Covid 19 pandemic were also the reason that the data collection took part over two time periods. The first data collection took place in May 2021 and the second data collection took place in September 2021. Both sessions followed identical procedures. All changes between the two blocks were documented in great detail. Those changes included construction work, new graffiti on a wall and different advertisement posters.

Participants were invited to take part in a free exploration walk and were all informed about the procedure of the eye-tracking session and the following questionnaires. Both testing phases went on from Monday until Sunday, always in the morning between 7 am and 10 am. Early mornings were chosen to prevent the participants from being distracted by open stores and to maintain light conditions. First, participants were greeted and verbally informed about the procedure of the eye-tracking session, as well as the short follow up questionnaire directly afterwards and a longer follow up questionnaire two weeks later. Participants received a QR-code to fill in the consent form and the short follow up questionnaire. After filling in the consent form, participants were equipped with the Pupil Labs Headset, called Pupil core (Figure 1).

The headset was connected to a portable hardware, a Microsoft Surface Pro 4. Participants carried the mobile hardware by themselves. Calibration was performed using the CalibMe software, while standing at a line 2-3m in front of the calibration marker (Santini, Fuhl, &

Kasneci, 2017). CalibMe is a new approach that stands for calibration with movements, which enables researchers to calibrate the mobile eye-tracking glasses quickly and precisely. Participants were instructed to focus on the centre of the calibration marker and to slowly rotate their head while fixating the centre. To avoid problems with the mobile eye-tracking device and direct sunlight, all participants were instructed to wear a baseball cap.

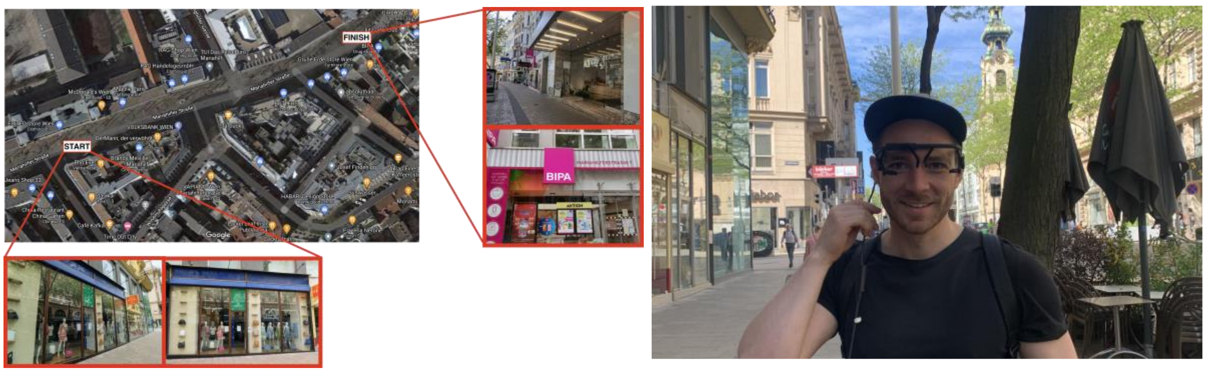


Figure 1. (A) Google maps satellite picture, showing the eye-tracking path at Mariahilferstrasse in Vienna. The small, embedded pictures show the start and ending point of the test path. (B) A photo of one of the participants with the Pupil labs headset**.**

After a successful calibration, all participants started from the same starting point at Mariahilferstrasse 33 (see Figure 1. A). They were instructed to walk down the street until they reached another meeting point 250m later at the BIPA store. At that point, one of the experimenters stopped the recording and started it again before participants walked back up to the initial starting point. As a free viewing task was conducted, the participants were not given any specific task to perform but were encouraged to explore the visual scene by themselves. They were instructed to walk straight ahead on the right side of the street and were informed that they were allowed to move their head without any restrictions. The instructions were chosen to be as unconstrained and natural as possible, but to still ensure that all participants would walk the same path. In case they needed any help or felt any discomfort, participants were asked to raise their hand. The participants were constantly followed by one of the experimenters during the walk at a distance of ten meters. This was followed by a short open recall task (see Appendix).

The output of the mobile eye-tracking is a video recorded by the first world camera with an estimated gaze point recorded by the two pupil cameras. For the evaluation of the eye-tracking data, categories of objects in the scene that were fixated frequently were predefined (see Table 1). Texts, but also other objects, captured attention and the following categories were formulated: Architecture (e.g. balconies, buildings), advertisements, nature (e.g., trees, green spots), people, traffic, street signs, shop signs, other text (e.g., graffiti, stickers, posters), and graphic elements (including the logos of the shops). Manual annotations were used for the evaluation of the videos, by assigning labels to the different objects with the eye movement coder (version 2.4). For the annotation process, the path of the video was chosen first (up/down), followed by the focused object out of a predefined list (see Table 1). The output of the eye-tracking data included the frames and the viewing time per object. Multiple looks were summed up and compared with other categories.

Table 1.

Different categories with definition

Category	Definition
People	Man, Woman, Child
Nature	Trees, Grass, Bushes, Potted Plants
Architecture	Balcony, Ornate Facades
Street signs	Mariahilferstrasse, Theobaldgasse, Königsklosergasse, Capistrangasse
Shop signs	All shop signs
Traffic	Cars, Bikes, E-scooters, Scooters
Advertisement	Posters, Billboards, Special offers
Other text	Stickers, Graffiti, Doctor Signs
Graphic elements	Logos, ATM sign

2.4 Photos

In the follow up questionnaire, the participants were presented with 40 pictures in a randomised order, 20 pictures of shop signs that were present at the path on Mariahilferstrasse that participants walked during the study and 20 photos of text signs that were not present and taken from other areas of the city as distractors. Since the signs were of different sizes, the distance to take the photo was chosen individually. All pictures were taken with the Huawei Mate 40 Pro and cropped to the same format of 19.5:5 before being included in the questionnaire (see Appendix).

2.5 Follow up questionnaire (online)

• Memory: Recognition task

• Evaluation: Rating task

• Language History Questionnaire (Li et al., 2020)

Additionally to the field study, a questionnaire was conducted. It included a memory task as well as different rating tasks. The results of the questionnaire were compared with the individual videos of each participant. The questionnaires were created on Survey Monkey. Due to the Covid 19 pandemic, participants completed the questionnaire on their own computer with help of a link.

Relatively little work has been conducted in a complex visual environment to understand how vision, aesthetics and memory are coordinated. For this reason the first questionnaire started with a recall task in form of open questions about what the participants could remember from the different categories (see Table 1). The participants were then asked whether they recognised the signs from the eye-tracking path. For this they were presented with 40 images of shop signs – 20 of which were truly from the test path and 20 that were distractors. To assess the aesthetic experience, the author used 7-point Likert scales and let participants rate all 40 photos for beauty, attractiveness, likability, interest, meaningfulness, and emotional arousal:

(1) *Beauty*: how beautiful is the shop sign?

(2) *Attractiveness*: how attractive is the shop sign?

(3) *Likability:* how well is the shop sign liked?

(4) *Interest*: how well the shop sign evoked their interest?

(5) *Meaning*: how meaningful is the sign for the participant?

(6) *Emotion*: whether the shop sign awoke emotions?

(7) *Familiarity:* whether they knew the shop sign?

Beauty, attractiveness, and likability ratings were chosen to investigate whether a shop sign attracts attention because of its appearance. Meaningfulness, emotional arousal, interest, and familiarity would give information whether a sign attracted attention for more than its optics. All ratings were given on a seven-point scale from 1 (fully agree) to 7 (fully disagree). Interest, liking and meaningfulness has already been used in studies concentrating on aesthetic experiences (Leder et al., 2006).

LHQ was included mainly to ensure that participants have similar reading backgrounds.

3. Results

First, the data of the 24 participants whose video material was complete was applied for more in-depth analysis. Here, a focus was placed on the general gaze distribution and the viewing time at different categories and shop signs. For the analysis of aesthetic rating and memorability of shop signs, the data of all 38 participants who completed the questionnaire was used. Finally, the results of the individual questionnaires were connected to the individual eye-tracking data. Therefore, the data of the 24 participants with complete videos and questionnaires was used. All statistical data analysis was conducted in R.

3.1 Eye tracking data

By inspecting the eye tracking data, the researchers could detect what participants looked at while walking along the test path at Mariahilferstrasse. In the analysis of the eye tracking data, focus was placed on how the attention of the participants was distributed between the different categories while walking along a commercial shopping street, as well as the memorability, the aesthetic evaluation, and the viewing time of the different shop signs. The participants were also questioned about the usability of the Pupil labs headset as part of the questionnaire. The comfort of wearing the eye tracking headset was assessed with a 5-Point Likert scale (1 = "not at all", 5 = "very comfortable"). The majority answered that they found it very comfortable with 44.44% followed by the second highest rating with 33.33% and the third highest with 22.22%. None of the participants chose the two lowest ratings, stating that the Pupil labs headset was uncomfortable to wear.

Eye-tracking data consists of the first world video recorded by the front camera in the headset and the frame-by-frame recordings from the estimated gaze position, recorded by two pupil cameras. To determine where participants looked at, each position was coded separately, manually in categories: people, traffic, architecture, nature, street signs, shop signs, advertisement, graphic elements and other text (see Table 1). The eye movement coder (version 2.4) was used for the manual annotations of each video. Nine different categories were selected for the manual annotations. Every detected fixation was categorised according to the predefined category list. The data collected from the mobile eye-tracking glasses were exported in a CSV file. When evaluating the data of the 24 participants, the researchers concentrated on the fixation locations and times of each participant individually in order to determine the individual differences in the gaze distribution.

Here the data of the 24 participants, were videos and questionnaires were complete were used. The total time of the annotated videos was 1 h 57 min 20 sec, on average 02:26 min for the way down and 02:28 min for the way up. This was counted from the time of the first to the last annotation of each video. This means that in total participants spent on average more time for the way up compared to the way down. This could be due to a slight upward slope.

3.1.1 Gaze distribution: different categories and viewing time

Participants were confronted with different objects during the walk at Mariahilferstrasse. To find out which objects captured gaze frequently, the total amount of the viewing time each participant spent on the different categories was compared (see Figure 2). The focus was placed on which categories would attract the most attention in a public shopping street. Comparing the viewing time towards these different objects, it could be shown that some drew more attention than others. Each fixation was coded in line with the predefined list of categories (see Table 2).

It can be seen that participants spent the most time, with 272.30 sec or 21.34 % of their total measured viewing time, looking at other people. This was followed by the mean viewing time at nature with 265.47 sec or 20.81 % of the total measured viewing time. The total time spent looking at shop signs was 234.47 sec or 18.38%.

The total mean viewing time combining all categories that included text – shop signs, street signs and other text - was 310.90 sec or 24.37 %. This is particularly interesting, since participants were not instructed to concentrate on text. Moreover, these results are in line with previous findings of the study of Cerf et al. (2009) and Wang and Pomplun (2012). The eye tracking data from this study shows a clear lack of interest in architecture with 60.70 sec or 4.76 % compared to the other categories.

Table 2.

Sum of viewing time in seconds for the different categories

	%	Sec
Nature	20.81	265.47
Traffic	14.88	189.83
People	21.34	272.30
Other Text	5.42	69.17
Advertisements	10.30	131.37
Graphic elements	3.54	45.20
Architecture	4.76	60.70
Shop signs	18.38	234.47
Street signs	0.57	7.27

It should be emphasised when comparing the individual answers, that a large variance between the participants can be observed. Regarding the amount of time that they focused on text, the results show that some spent a large amount of time focusing on shop signs compared to the other categories with more than 50% of their viewing time and others almost did not look at them at all with only 1% of their total viewing time (Figure 2).

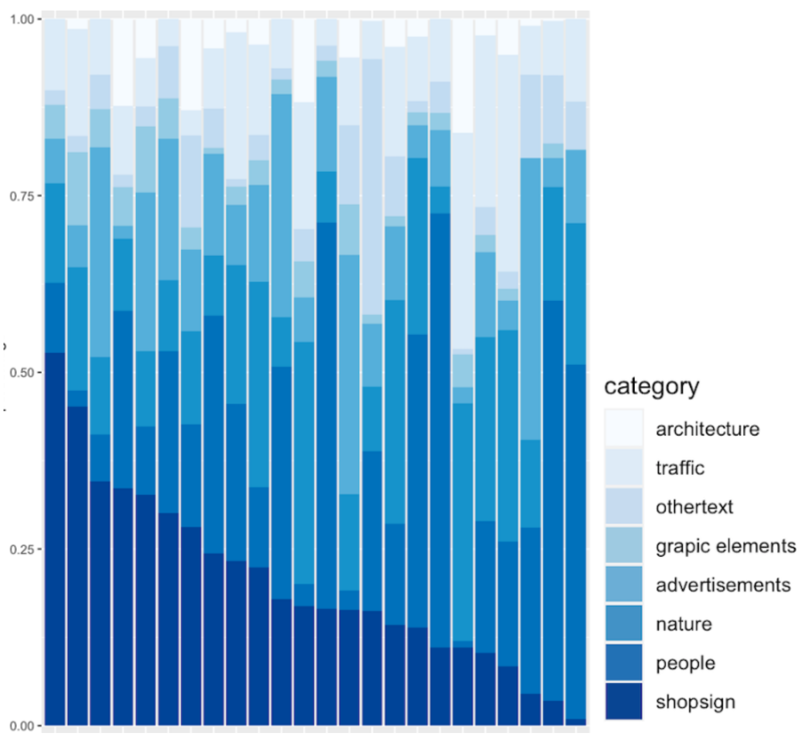


Figure 2. Individual viewing time sorted by highest viewing time of shop signs

3.1.2 Viewing time of different shop signs

To understand how participants explore and engage with text signs, focus was placed on the shop signs that participants encountered while taking part in the study. Analysing the time participants spent looking at shop signs, it became obvious that participants spent varying amounts of time looking at the different shop signs. Table 3 shows an overview of the statistical analysis for the viewing time of the different signs. The mean fixation duration showed major differences between the different shop signs, with the longest mean viewing time being for "Interio", followed by "Phoever". Compared to the longer viewing times for some of the signs, there were two shop signs that were looked at either very shortly or not at all: "Öfferl" and "To go". As such, comparing the viewing times of the shop signs, it could be shown that there was a major difference in the attentional distribution.

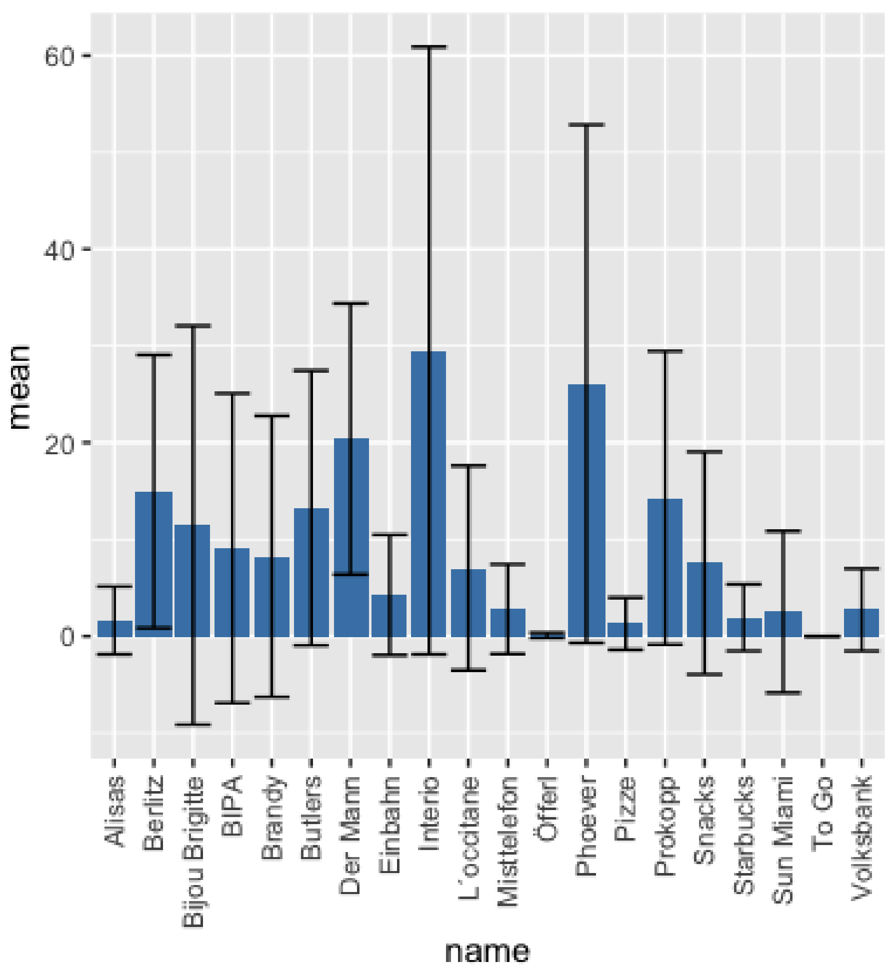


Figure 3. Mean and SD of viewing time on the different shop signs

Table 3.

Viewing times in seconds for the different shop signs

Shop Sign	Mean	SD	Median	IQR
Interio	29.50	31.39	18.0	29.75
Phoever	26.08	26.77	18.0	29.50
Der Mann	20.37	14.01	17.5	14.25
Berlitz	14.96	14.10	12.5	19.75
Prokopp	14.29	15.16	11.0	12.50
Butlers	13.25	14.20	9.0	19.00
Bijou Brigitte	11.46	20.61	4.0	7.25
BIPA	9.08	15.99	4.0	10.00
Brandy	8.25	14.53	3.0	7.50
Snacks	7.54	11.49	3.0	5.75
L´Occitane	7.04	10.57	2.0	10.25
Einbahn	4.25	6.22	2.0	6.00
Misttelefon	2.79	4.63	1.0	3.50
Volksbank	2.71	4.25	2.0	3.00
Sun Miami	2.50	8.35	0.0	0.00
Starbucks	1.92	3.45	0.0	2.00
Alisas	1.63	3.51	0.0	1.25
Pizze	1.29	2.64	0.0	1.25
Öfferl	0.08	0.28	0.0	0.00
To Go	0.00	0.0	0.0	0.00

Figure 3 shows that some shop signs e.g. “Interior”, “Phoever”, and “Der Mann” had a higher mean viewing time than others. Interestingly the standard deviation is the largest for the two signs with the highest mean viewing time.

Table 4.

Number of participants that looked at the shop sign

Shop Sign	Number of people that looked at the sign
Der Mann	24
Interior	23
Berlitz	23
Butlers	22
Prokopp	22
Bijou Brigitte	21
Phoever	21
Bipa	20
Brandy Melville	18
Snacks	18
L’Occitane	17
Einbahn	15
Volksbank	15
Misttelefon	15
Starbucks	11
Pizze	8
Alisas	8
Sun Miami	3
Öfferl	2
To Go	0

Table 4 shows the amount of participants that had actually looked at the individual signs. What can be taken from this, is that even though “Interio” had the highest mean viewing time, it is only in second place with regards to the number of people to look at it. On the other hand, “Der Mann” was fixated on by every single one, but it wasn’t done so for the longest time. This shows that not only viewing time is of importance but also the variance.

3.2 Aesthetic rating and memorability of shop signs

3.2.1 Memorability of shop signs

In this part of the analysis all participants that took part in the study were included, also the ones where the eye-tracking data could not be used. All of the 39 participants completed the follow up questionnaire between one and four weeks after taking part in the field experiment (M = 13.5 days, SD = 7.31). Memory of the shop signs was assessed with the help of a recognition task and an open recall task in the follow up questionnaire (see Appendix). In the recognition task, the participants were asked if they remembered a specific shop sign from the path that they walked along Mariahilferstrasse while taking part in the study, requiring a yes or no response. The purpose of this was to find out how well participants remembered the different text signs that they encountered during their walk. Analyses revealed that from the 39 participants 45.24% recognised the pictures of the shop signs that were presented correctly (see Figure 4). In contrast, they rejected the distractors – the shop signs that could not be recognised by the participants as having been seen in the study - correctly with 91.85% (see Figure 4). The pictures from the study path received a higher number of positive recognition ("yes") replies in total, compared to the distractor shop signs. Overall, it can be said that there is better recognition in form of correct rejection for the shop signs that were not present in the path.

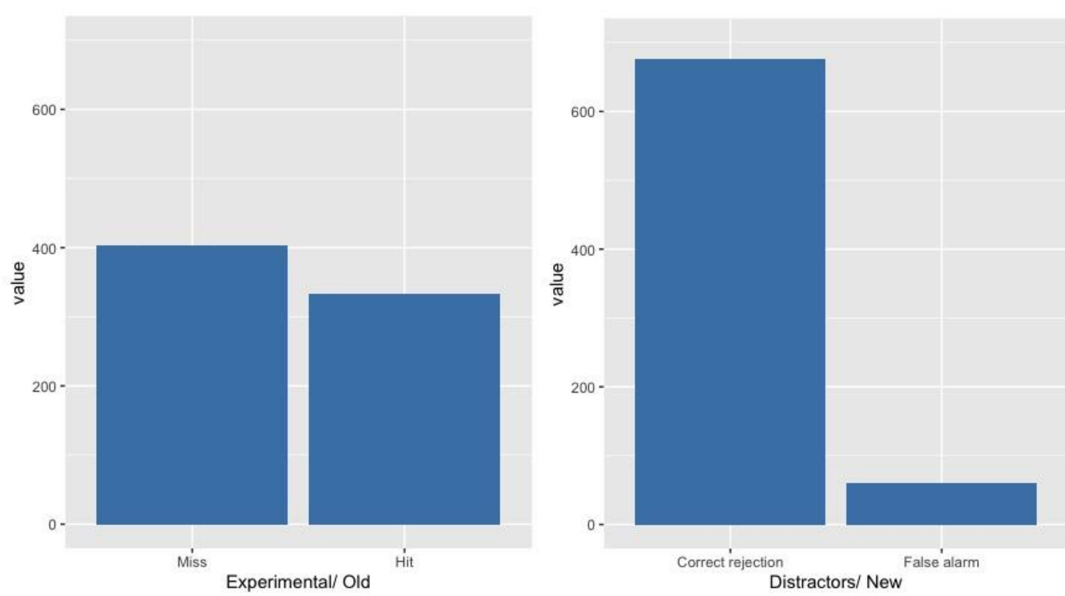


Figure 4. Results recognition task

Later on, the individual ratings for each sign that was present in the study will be compared specifically with the individual eye-tracking data from each participant. For this only the data of the 24 participants will be used, whose eye-tracking data was complete.

3.2.2 Aesthetic rating of shop signs

Next to the recognition task, the researchers also conducted an aesthetical evaluation for all the shop signs (see Appendix). The aesthetical evaluation was chosen to show whether there would be differences between the individual signs and later on to also investigate what role the aesthetic evaluation plays in directing the viewing behaviour. Therefore, the pictures of the 40 shop signs (20 from the walk, 20 distractors) were presented again and the participants were asked to evaluate the pictures with the help of a 7-point Likert scale. The rating included questions for beauty, attractiveness, likability, interest, meaningfulness, emotional arousal and familiarity.

To present the results in a more comprehensive way, a matrix was constructed to show the connection between the different ratings of the 40 photos (see Figure 3). Notably, the familiarity of the signs was not highly correlated to the aesthetic evaluations. Other than that, there is a high positive correlation between the different evaluation ratings for text. The highest correlation can be observed between attractiveness and meaningfulness with 0.9. The lowest correlation on the contrary, can be observed between interest and familiarity, with a positive correlation of 0.34.

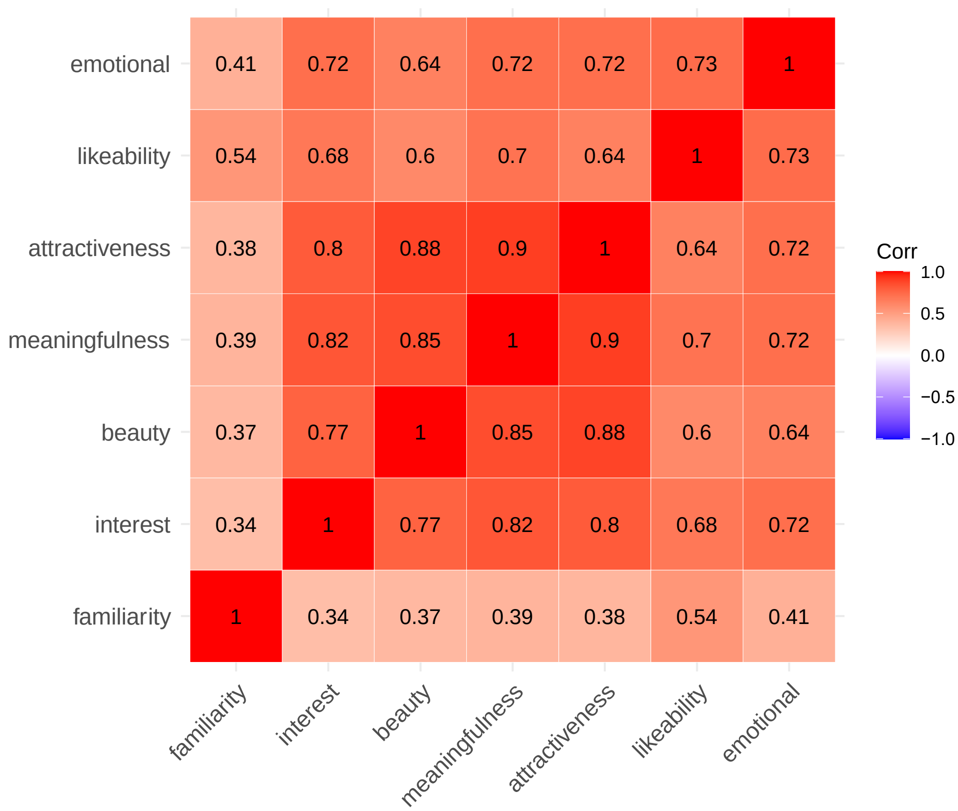


Figure 5. Correlation matrix between the different ratings

Additionally, focus was placed on the beauty evaluation of the different shop signs (see Figure 6). The results show a high variation in how beautiful the participants evaluated the different shop signs. The standard deviation shows that the variance is mostly low, so there is not much noise observed around the mean. In some cases however, the mean ratings were influenced by larger variance. Mean liking was the highest for the “L’Occitane” shop sign with 5.46. The lowest mean liking was observed for the “Prokopp” sign with 1.56. See Figure 6 for a chart and a plot for a better overview over the different beauty evaluations of the shop signs.

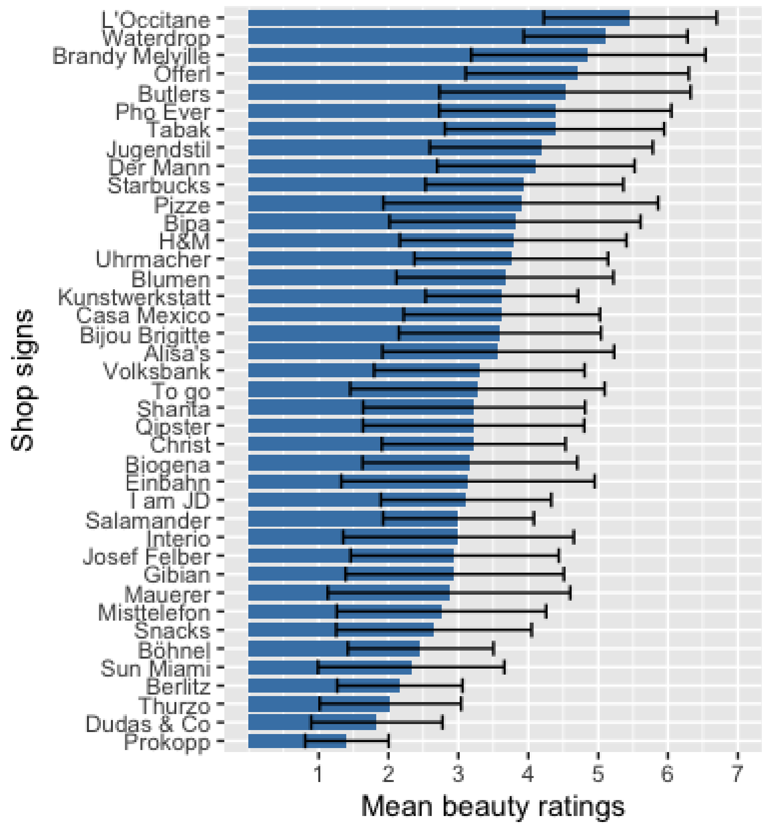


Figure 6. Mean rating beauty and SD of all shop signs

Furthermore, the influence of size, position and colour of the different shop signs was investigated. Concentrating on the size the results show a positive correlation, that bigger signs caught more attention compared to smaller sized ones, $r = .39$, $p < 2.2e\text{-}16$. Also, for the position a positive correlation was observed, that signs that were on a higher position had a higher viewing time compared to signs in a lower position, $r = .26$, $p = 1.212e\text{-}08$ Finally, the results also show a positive correlation for the colour of the signs, meaning that colourful signs had a higher viewing time compared to signs in black and white, $r = .26$, $p = 1.391e\text{-}08$.

3.3 Memorability and viewing behaviour for shop signs

To investigate the influence of the viewing behaviour – whether participants really looked at a shop sign or not - on the performance in the recognition task, the individual ratings for each sign that was present in the study are compared to the individual eye-tracking data of each participant. The data of the 24 participants with complete videos of the eye-tracking path and their individual questionnaires was used. Here, the data only includes the 20 shop signs that were present in the study path without the distractor shop signs.

The results show that the total amount of hits (i.e. participants looking at the signs and then recognising them correctly in the questionnaire) is 30,68 %. The misses (i.e. participants not looking at the sign and not recognising it in the questionnaire) is 15.91%. The false positive score (i.e. participants responded in the questionnaire that they had seen the sign, when in fact they hadn’t looked at it) is the lowest with 9.09%. The false negative score (i.e. participants had not looked at the sign but reported that they had seen it) was the highest with 44.32%.

3.3.1 Aesthetic value and viewing behaviour

To examine which role the aesthetic evaluation of the shop signs has on the viewing time, a product moment correlation was used. Here all the individual rating scores for beauty, attractiveness, likability, interest, meaningfulness, emotional arousal, and familiarity of each participant are compared with their individual viewing time duration. By connecting the eye-tracking data of the individual videos to the according questionnaires, the data indicates that some individual participants show a significant effect (see Appendix). Examining the data of all 24 participants, no connection between the different aesthetic evaluations and the viewing time of the shop signs could be shown. There seems to be no correlation between the individual ratings of the signs and the time participants spent looking at the signs (see Table 5).

Table 5.

Pearson's product-moment correlation

	Correlation	KI´s	P- Value
Beauty	-0.01013476	[-0.11, 0.08]	0.84
Attractiveness	0.002159426	[-0.09, 0.09]	0.96
Likeability	-0.002823105	[-0.09, 0.09]	0.95
Interest	0.007510817	[-0.08, 0.10]	0.88

Meaningfulness	-0.004142818	[-0.10, 0.09]	0.93
Emotional	0.03224394	[-0.06, 0.12]	0.51
Familiarity	-0.001556782	[-0.09, 0.09]	0.97

3.4 Language History Questionnaire

The Language History Questionnaire (Li et al., 2020) was given to ensure that the participants had similar reading experiences and language backgrounds. It is an important tool to assess the language background. All participants stated that their current main residence is Austria. The analysis of the data shows that no participant had to be excluded. All participants showed similar language and reading backgrounds. They all had English lessons in school and 37,5% participated in an language exchange program. In this study only native German speakers were involved but as this study is a pre-study for a larger study that will take place in the spring of 2022 it was already included in this questionnaire.

4. Discussion

This thesis examined, in an exploratory fashion, where people look at and how they engage and explore text in an everyday life setting. Real world scenes contain a large amount of information, but humans are limited in the capacity of what they can absorb and process. In their daily lives they are confronted with all different kinds of texts. First, the gaze distribution between the different categories was investigated, followed by a more in depth analysis of the exploration of text in a public space. The memorability and aesthetic value of shop signs was examined, and attention was paid to the relationship between viewing time with memorability and the aesthetic evaluation of text. Visual attention is a selective process and with the help of mobile eye-tracking it could be shown where participants looked at while walking in a natural environment. Walking through one of Vienna's biggest shopping streets was utilised as an aesthetic experience. Eye-tracking only records the eye movements, therefore additional questionnaires were used to interpret the recordings. In a follow up questionnaire, participants completed a memory recognition task and an evaluation rating task.

4.1 Fixation location: Which objects capture gaze frequently?

Regarding the first research question, which objects capture gaze frequently, the data indicated that other pedestrians attract the most attention while walking through a shopping street, followed by nature. Researchers could already show that faces have an immense biological and social importance. Bindemann et al. (2015) found that faces attract visual attention to the detriment of other visual stimuli, by comparing faces directly with other classes of stimuli. Even infants detect human faces in complex visual environments (Kelly et al., 2019). Shimojo et al. (2003) could show that participants looked longer at isolated faces that they found to be more attractive. Leder et al. (2016) replicated the positive relation between looking time and beauty with photographs of natural scenes that always pictured two people. The results of the present study are in line with these results that people attract attention.

The large amount of time spent looking at nature is particularly interesting, since the testing took place in a truly urban setting. Researchers could already demonstrate that people prefer nature scenes over human made scenes (Biederman & Vessel, 2006; Tinio & Leder, 2009).

Van Den Berg et al (2007) pointed out that due to the high settlement density in cities looking at nature can help psychological restoration.

When combining the different kinds of text (i.e., shop signs, street signs and other text) the results show that text draws the most attention compared to all other categories. It stands out, as the participants were not instructed to concentrate on text. Wang and Pomplun (2012) already suggested that during a free viewing task in the laboratory, observers might attempt to receive as much information as possible and therefore concentrate on text to make a scene more interpretable and contribute to its understanding and memorization. The results are in accordance with the results of Cerf et al. (2009); whereby people fixate on faces and text more often, compared to other objects. Concentrating on the individual patterns between the participants, high variation was observed. While some participants for example, spent more than 50% of their overall viewing time on text signs, others only concentrated 1% of their overall viewing time on them. The results of the present study extend the understanding of gaze control in a natural environment.

One limitation of this study is consequently linked to the study design and the use of a field study. Eye-tracking in real life has an implicit difficulty in the control of variables to guarantee that all participants experience the same conditions. Since an exploratory design was used for this study, it was neither possible nor desirable to control those. On a shopping street, one will almost always encounter people, as well as other distractors, such as traffic. The aim was to allow natural eye movements without any restrictions. It could happen that unexpectedly appearing stimuli, for example, a pedestrian who comes around a corner, would draw a participant’s attention while walking down Mariahilferstrasse. For future research in this field, the researchers suggest combining controlled laboratory sessions with real world field studies.

4.2 How do people engage and explore text in an everyday life setting?

The analysis of the eye-tracking data of the 24 participants indicate that participants spent the highest viewing time on text elements that they encountered during their walk compared to other categories. Regarding the different shop signs, the data shows large variations between the mean viewing time spent on the different signs. The highest mean viewing time was observed for the “Interio” shop sign, followed by the “Phoever” shop sign. On the other hand, other shop signs didn’t get any or only little attention. Comparing the appearance of the shop signs in form of size, position and colour, significant results were observed. It could be shown, that larger signs caught more attention compared to smaller sized ones. Moreover, a positive correlation for the position of the shop signs was observed, whereby signs that were placed in a higher position had a longer viewing time compared to signs in a lower position. Wang and Pomplun (2012) could already demonstrate that typical locations of text attract the attention of participants even if there is no text there. Finally, the results in the present study indicated that coloured signs attracted more attention than black and white signs.

Surprisingly, some international and well-known shop signs, such as the Starbucks sign, did not catch the attention of the participants as much as expected. Even without actually having looked at the written sign, participants were still able to correctly identify its presence on the path in the follow up questionnaire. A possible reason could be that participants focused on other elements, such as the logo of Starbucks, that was categorised as a graphical element instead. Mollerup (2005) pointed out the distinction between ‘sign’ and ‘SIGN’ to differentiate between the everyday sign and the semiotic SIGN respectively. Harland (2015) created a model that showed the text sign in the middle and the semiotic sign as a circle around it, which includes the logo or the distinguishing colour that gives a wider picture (see Figure 7). In the case of the Starbucks sign, this would mean that a person could look at the logo or even at the classical dark green colour of the awning and already associate that with Starbucks without looking at the text sign itself. This leads to the idea that reading is much more than text detection itself.

It should be noted that individuals will have differing interests and those will inform how they perceive texts in public spaces like the Mariahilferstrasse. Similarly, the way one moves will affect their gaze distribution and therefore two people walking the same path, will likely see texts differently. Additionally, all the different text forms lead to a high degree of differentiation in the gaze distribution.

Finally, it is important to mention that these results are limited to the location that was chosen, i.e., Mariahilferstrasse in Vienna. Future analysis could also compare different locations that contain text in more depth and combine laboratory and field settings; this would show whether results would repeat themselves. It should be remembered that unlike in a laboratory experiment, it is not possible to draw direct conclusions from a field study. In this case it did however, deliver important insights to understanding attention distribution in real-world scenes.

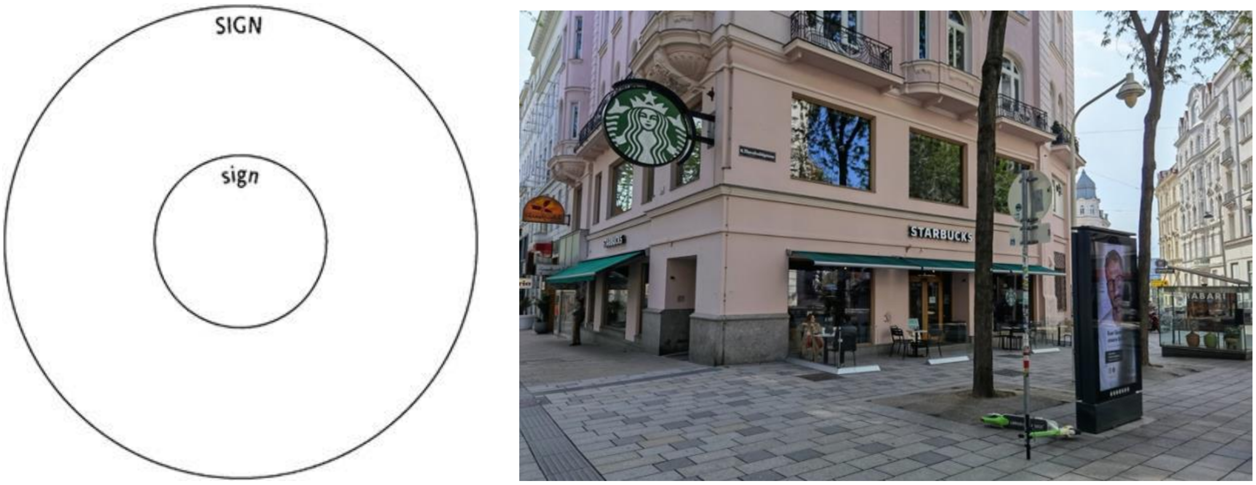


Figure 7. The everyday ‘sign’ and semiotic SIGN. Source: Mollerup (2005, 10), redrawn by Harland (2015). (B) Photo of the Starbucks store showing the text sign and graphical element.

4.3 Memorability and aesthetic evaluation

The results of the memorability and aesthetic evaluation of text give important insights into how participants perceive their environment. Regarding the recognition task of the 40 pictures of the shop signs – including the 20 pictures from the test path and the 20 distractors - the results show that participants recognised 45.24 % correctly from the pictures of the shop signs that were present while taking part in the experiment. They correctly rejected the distractors in 91.85% of the cases. As said before these results indicate that there is better recognition in form of correct rejection for the shop signs that were not present in the test path. Moreover, participants showed better memory recognition, by answering yes that they saw the signs while taking part in the study for signs that were actually present.

Looking at the evaluation of how beautiful participants rated the different shop signs, the data shows a high variation. The “L’Occitane” shop sign had the highest mean beauty rating. It is one of the most colourful signs along that section of Mariahilferstrasse. In the present study a positive correlation between the colour and the viewing time was observed. Further studies should be conducted to also concentrate whether there is a connection between the different colours of the shop signs and the performance in memorability. Dzulkifli and Mustafar (2013) could already show, that colour has a potential to increase chances of environmental stimuli to be encoded, stored, and retrieved successfully.

Text in real life is a rich area for future research. A shopping street especially, is filled with multi orientated texts. It would be interesting to compare outstanding protruding shop signs with fascia shop signs (wall mounted signs) in future studies. Kim and Park (2020) found that high signboards are fixated on more often, compared to low signboards. There was also one rotating sign in the test path, which attracted a lot of attention and was mentioned frequently in the open questions (see Appendix).

4.4 Comparing of the eye tracking data and the memorability and aesthetic evaluation

Noton and Stark (1971) introduced the "Scanpath theory" which predicts that repeated eye fixations will help observers recognise a previously seen image. Foulsham and Kingston (2013) found that the fixation during the encoding of images was a predictor for recognition accuracy.

Therefore the present study also compared the mobile eye-tracking data with the individual answers of the questionnaires of the 24 participants with the complete data, to figure out whether there is a connection between the fixation time and location and the answers in the questionnaire. Here also task dependency should be considered. Droll and Eckstein (2009) suggested that the task is important for the performance of participants. They demonstrated that participants expecting to receive a memory task had longer fixation duration on the experimental objects, while walking outside around a building, compared to observers simply instructed to walk as they would normally. In this study, the experimenters exchanged some of the objects and they could show that participants who were expecting a memory task had longer fixation and a higher chance to detect whether those objects had been exchanged. Here, in the present study, a free exploration task was chosen to allow participants to interact freely with their environment and to ensure a natural and realistic experience for the participants while taking part in the study.

Little is known about gaze behaviours and memory in a real-world environment. In order to answer the third and fourth research question, whether the aesthetic value has an impact on the viewing behaviour or on the memorability of text, the individual answers in the questionnaires were compared with the total viewing time of each individual participant. By connecting the individual eye-tracking videos with the questionnaires, no significant results could be shown between the memorability and viewing time. These results are limited by the fact that there is a high variation in time passed between the walk and the follow up questionnaire. These time differences vary between one and four weeks (Mean= 13.5 days, SD= 7.31). Notably, there is no significant correlation found between the time that passed until participants answered the follow-up questionnaire and their performance in the recognition task (r = 0.14 , p = 0.55). Even though there is no correlation between the time passed and the performance in the recognition task, it is recommended for future studies to let participants fill out the recognition task directly after the walk and again two weeks later to compare the different results. Interestingly, when concentrating on the individual performance, the results show great differences between the participants. Some performed much better in the recognition task and recognised the shop signs in 85% of the cases correctly compared to other participants who only recognised 5% of the present shop signs (see Appendix).

Regarding the aesthetic evaluation task, there seems to be no correlation between the aesthetic evaluation of the shop signs and the time participants spent looking at them. This could be caused by the different visibility of the signs, depending on the position. Some attracted more attention due to their prominent position compared to others. However, it is important to note when concentrating on the individual answers, the results indicate high differences between the different participants. Some participants showed a significant correlation between the aesthetic evaluation and the viewing time (see Appendix). It is noteworthy, since it points out that people have different experiences while walking in the same public space. These results indicate that the aesthetic value of the text signs can predict the gaze distribution and viewing time.

By comparing the eye-tracking data with the individual questionnaires from the 24 participants with usable complete videos and questionnaires, the data shows that participants looked at the signs and recognised them correctly in the questionnaire in 30.68% of the cases. That participants didn't look at the sign but reported that they had seen it, occurred in 44.32% of the cases. One of the main challenges of mobile eye-tracking is that it only allows researchers to conduct the raw gaze behaviour in the form of fixations. Eye-tracking only records and displays ocular fixations but cannot measure the peripheral vision which makes up 98% of the visual field (Land, 1999, cite after Rose, 2015). It is important to mention that ocular fixations do not represent everything that has been seen by the participants, but the area fixated in that moment. Peripheral vision allows the participant to choose where to fixate on next. It is also possible to fixate on some specific visual elements, but move the attention to something else or, to put it in other words, fixate on something without paying direct attention to it. It could also be seen that participants falsely recognised shop signs in 9.09% of the cases in the recognition task, without showing any direct fixation of that specific sign in the eye tracking data.

4.5 Limitations

The manual coding of the eye tracking data was a very lengthy process, so for future research these time-consuming limitations could be overcome with the help of a machine learning program. These would annotate the videos automatically once the researchers had identified relevant objects. Another recommendation for further studies that use eye-tracking in real life would be to have a testing path extending over the area of interest so that it would remain pure of distractions, since for this study the researchers could always be seen at the beginning and at the end of the recorded videos.

The Language History Questionnaire showed that all participants were native German speakers with similar reading skills. Viennese public spaces, especially the shopping streets, are filled with English language texts. Therefore, for future research it would also be interesting to test non-German speaking people and compare the data based on signs in different languages.

Another limitations is the relatively young age of the eye-tracking participants with a mean age of 23.7 (SD=3.2). This could be a potential confounder considering their interest and viewing behaviour. Therefore, a more diverse population regarding age and gender would be recommended.

The development of affordable portable eye-tracking headsets is still relatively recent, therefore there is only a limited number of studies conducted in natural, everyday settings. The ongoing development of technologies offer increasing opportunities to the use of mobile eye-tracking. It is clear, that there is great potential for future studies.

5. Conclusion

This thesis investigated the connection between memorability, aesthetic evaluation and viewing time of texts that can be encountered in everyday life. A mixed method design which included eye-tracking and multiple types of questionnaires was employed. The free exploration task allowed the participants to interact without restrictions with their environment.

The experimental results of the eye-tracking data indicate that text draws attention in a commercial shopping street. The data of the mobile eye-tracking shows that text, with mean viewing time of 22.77%, draws the most attention, followed by people and nature. These results are relevant, since participants were not instructed to look at text specifically. The study supports the findings of the studies of Cerf et al (2009) and Wang and Pomplun (2012).

With regards to memorability the results showed that participants were better able to identify signs they hadn’t seen than those they had, which could be an interesting topic for future research. Furthermore, it could be shown, even without actually reading a shop sign, participants could still recognise a store which leads to the conclusion that information can be gathered through not only text detection but also the recognising of graphical elements, colours and other objects.

In addition to these important insights into the understanding of gaze distribution in a real world environment, the present study also has highlighted important implications for the use of mobile eye-tracking glasses outside of the laboratory. The limitations of this use have been discussed in great detail, however it does offer a high potential for the application in future studies.

The study presented in this thesis would profit greatly from being repeated on a larger scale bearing in mind the lessons that have been learned in this smaller scale. Such studies could give useful information to be applied not only in research but also in the planning of public spaces.

References

Bindemann, M., Burton, A., Hooge, I., Jenkins, R., & de Haan, E. (2005). Faces retain attention. *Psychonomic Bulletin & Review*, 12, 1048–1053.

Biederman, I., & Vessel, E. (2006). Perceptual Pleasure and the Brain. *American Scientist*, *94*(3), 247. https://doi.org/10.1511/2006.59.247

Brieber, D., Nadal, M., & Leder, H. (2015). In the white cube: Museum context enhances the valuation and memory of art. *Acta Psychologica, 154*, 36–42. doi:10.1016/j.actpsy.2014.11.004

Brieber, D., Nadal, M., Leder, H., & Rosenberg, R. (2014). Art in Time and Space: Context Modulates the Relation between Art Experience and Viewing Time. *Plos ONE*, *9*(6), e99019. https://doi.org/10.1371/journal.pone.0099019

Brockmole, J. (2009). *The visual world in memory*. Taylor & Francis.

Cashon, C., & Cohen, L. (2003). The construction, deconstruction, and reconstruction of infant face perception. In O. Pascalis & A. Slater (Eds.), *The development of face processing in infancy and early childhood: Current perspectives* (pp. 55–68). New York: NOVA Science Publishers.

Cerf, M., Frady, E., & Koch, C. (2009). Faces and text attract gaze independent of the task: Experimental data and computer model. *Journal Of Vision*, *9*(12), 10-10. doi: 10.1167/9.12.10

Cristino, F., & Baddeley, R. (2009). The nature of the visual representations involved in eye movements when walking down the street. *Visual Cognition*, *17*(6-7), 880-903. https://doi.org/10.1080/13506280902834696

Droll, J., & Eckstein, M. (2009). Gaze control and memory for objects while walking in a real world environment. *Visual Cognition*, *17*(6-7), 1159-1184. https://doi.org/10.1080/13506280902797125

Dzulkifli, M. A., & Mustafar, M. F. (2013). The influence of colour on memory performance: a review. *The Malaysian journal of medical sciences : MJMS*, *20*(2), 3–9.

Einhäuser, W., Kruse, W., Hoffmann, K., & Ko¨nig, P. (2006). Differences of monkey and human overt attention under natural conditions. *Vision Research*, 46, 1194–1209.

Fechner, G. T. (1876). *Vorschule der Aesthetik*. Breitkopf & Härtel.

Foulsham, T., & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *Journal Of Experimental Psychology: General*, *142*(1), 41-56. https://doi.org/10.1037/a0028227

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, *51*(17), 1920-1931. https://doi.org/10.1016/j.visres.2011.07.002

Harland, R. (2015). Graphic Objects and their Contribution to the Image of the City. *Journal Of Urban Design*, *20*(3), 367-392. https://doi.org/10.1080/13574809.2015.1031211

Henderson, J. (2003). Human gaze control during real-world scene perception. *Trends In Cognitive Sciences*, *7*(11), 498-504. https://doi.org/10.1016/j.tics.2003.09.006

Holmes, T., & Zanker, J. (2012). Using an Oculomotor Signature as an Indicator of Aesthetic Preference. *I-Perception*, *3*(7), 426-439. https://doi.org/10.1068/i0448aap

Honey, C., Kirchner, H., & VanRullen, R. (2008). Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *Journal of Vision*, 8(12):9,

Huerta, R. (2010). I Like Cities; Do You Like Letters? Introducing Urban Typography in Art Education. *International Journal Of Art & Design Education*, *29*(1), 72-81. https://doi.org/10.1111/j.1476-8070.2010.01633.x

Huey, E. (1898). Preliminary Experiments in the Physiology and Psychology of Reading. *The American Journal Of Psychology*, *9*(4), 575. doi: 10.2307/1412192

Irwin, D. E. and Zelinsky, G. J. 2002. Eye movements and scene perception: Memory for things observed. *Perception and Psychophysics*, 64(6): 882–895

Jaderberg, M., Simonyan, K., Vedaldi, A., & Zisserman, A. (2015). Reading Text in the Wild with Convolutional Neural Networks. *International Journal Of Computer Vision*, *116*(1), 1-20. https://doi.org/10.1007/s11263-015-0823-z

Kelly, D., Duarte, S., Meary, D., Bindemann, M., & Pascalis, O. (2019). Infants rapidly detect human faces in complex naturalistic visual scenes. *Developmental Science*, *22*(6). https://doi.org/10.1111/desc.12829

Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A Cognitive View of Reading Comprehension: Implications for Reading Difficulties. *Learning Disabilities Research & Practice*, *29*(1), 10-16. https://doi.org/10.1111/ldrp.12025

Land, M. F. and Hayhoe, M. 2001. In what ways do eye movements contribute to everyday activities?. *Vision Research*, 41: 3559–3565.

Land, M. F. and Tatler, B. W. 2001. Steering with the head: The visual strategy of a racing driver. *Current Biology*, 11: 1215–1220.

Leder, H., Mitrovic, A., & Goller, J. (2016). How Beauty Determines Gaze! Facial Attractiveness and Gaze Duration in Images of Real World Scenes. *I-Perception*, *7*(4), 204166951666435. https://doi.org/10.1177/2041669516664355

Leder, H., Tinio, P. P. L., Fuchs, I. M., & Bohrn, I. (2010). When attractiveness demands longer looks: the effects of situation and gender. *Quarterly Journal of Experimental Psychology*, 63(9), 1858–1871. doi:10.1080/17470211003605142

Lee, M., & Song, I. (2018). A Study on the Visual Continuity Expressed on Facade and Signboard of Commercial Street. *A Journal Of Brand Design Association Of Korea*, *16*(4), 183-194.

Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, *23*(5), 938-944.

Luck, S. J. and Vogel, E. K. 1997. The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657): 279–281.

Melcher, D. 2006. Accumulation and persistence of memory for natural scenes. *Journal of Vision*, 6(1): 8–17.

Mollerup, P. 2005. *Wayshowing*. Baden: Lars Müller Publishers.

Noton, D., & Stark, L. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, *11*(9), 929-IN8. https://doi.org/10.1016/0042-6989(71)90213-6

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42, 107–123.

Pearce, M., Zaidel, D., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics. *Perspectives On Psychological Science*, *11*(2), 265-279. https://doi.org/10.1177/1745691615621274

Posner, M., Snyder, C., & Davidson, B. (1980). Attention and the detection of signals. *Journal Of Experimental Psychology: General*, *109*(2), 160-174. https://doi.org/10.1037/0096-3445.109.2.160

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422. doi: 10.1037/0033-2909.124.3.372

Rayner, K. (2009). The thirty fifth Sir Frederick Bartlett Lecture: eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* 62, 1457–1506. doi: 10.1080/17470210902816461

Rosa, P. (2015). What do your eyes say? Bridging eye movements to consumer behavior. *International Journal Of Psychological Research*, *8*(2), 91-104. doi: 10.21500/20112084.1513

Tinio, P., & Leder, H. (2009). Natural scenes are indeed preferred, but image quality might have the last word. *Psychology Of Aesthetics, Creativity, And The Arts*, *3*(1), 52-56. https://doi.org/10.1037/a0014835

Santini, T., Fuhl, W., & Kasneci, E. (2017). CalibMe: Fast and unsupervised eye tracker calibration for gaze- based pervasive human-computer interaction. *Proceedings of CHI*, 2594–2605.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317-1322. https://doi.org/10.1038/nn1150

Underwood, G. (1998). *Eye guidance in reading and scene perception*. Elsevier.

Wang, H., & Pomplun, M. (2012). The attraction of visual attention to texts in real-world scenes. *Journal Of Vision*, *12*(6), 26-26. doi: 10.1167/12.6.26

Wheeler, M. E. and Treisman, A. 2002. Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131(1): 48–64

Image Index

Table Index

Appendix

Tabell 1. Pictures of the different shop signs

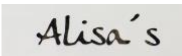
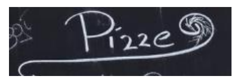
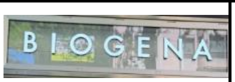
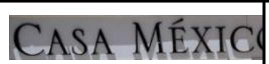
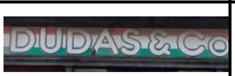
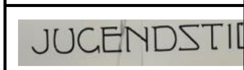
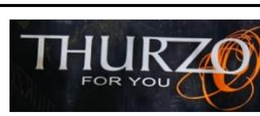
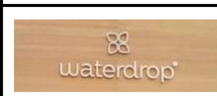
Eye-tracking				
				
				
				
				
Distractors				
				
				
				
				

Table 2. Percentage of the individual recognition of the present signs and the distractors

Participant	recognition present signs	False recognition distractors
1	0.85	0.05
2	0.25	0.00
3	0.60	0.10
4	0.50	0.00
5	0.60	0.50
6	0.40	0.00
7	0.65	0.00
8	0.65	0.00
9	0.15	0.00
10	0.60	0.15
11	0.35	0.00
12	0.60	0.20
13	0.55	0.20
14	0.50	0.30
15	0.65	0.00
16	0.45	0.05
17	0.60	0.15
18	0.55	0.20
19	0.50	0.12
20	0.25	0.05
21	0.50	0.05
22	0.45	0.00
23	0.15	0.00
24	0.05	0.05
25	0.40	0.00
26	0.40	0.10
27	0.25	0.00
28	0.50	0.00
29	0.45	0.00
30	0.40	0.00
31	0.35	0.15
32	0.20	0.00
33	0.50	0.05
34	0.50	0.25
35	0.45	0.20
36	0.45	0.05
37	0.50	0.05

Tabell 3. Answers open recall task

Shop Signs	Architecture	Street signs	People	Advertisements	Nature	Text
Albertina	Space between starbucks and volksbank (4)	Parking sign	Elderly gentleman, lady with dog	Exhibition poster for Albertina (4)	Trees (16)	Menus (4)
Audi	Shop windows	Construction site markings	People with dogs	Rotating sign (4)	Sun	Graffiti (1)
Bijou	Signs in facades	One way (2)	Woman with dog	Starbucks sign	Dogs (2)	Street signs
Brigitte (4)	Leiner House	Bike	Hardly but passed a few dogs	Kebab ads	(dog peeing on a pole)	Stickers
Bipa (16)	A nice modern house	Street Names (1)	People probably on way to work	Erste bank billboard	A lot of concrete	Information about organic products next to a bakery
Brandy Melville (7)	Old style	(Theobaldgasse (2)	Some people staring at me because of the glasses (2)	20% offer in Interio shop window	Grasses planted under some trees	Cut off lettering on glass door of Butlers
Butlers (6)	Pedestrian zone setting	Königsklostergasse)	employee of café	Audi		Ads on poles
Der Mann (4)	Stucco facades	Street Sign (2)	People standing in front of bakery	Raiffeisen		
Dm (2)	Wrought iron balconies	Mariahilferstrasse street sign (2)	Man with open black backpack	Bipa window advertising		
Hemp (2)	Construction site (2)		Yes but not in detail	Bakery offer on sign in front of shop		
Hotel sign (2)	Corner house w/ Bijou Brigitte			Ad billboard (2)		
Interio				In shop windows		
Le Pho (2)				Blackboard of daily specials for pizza		
L‘Occitane				Ad for new smartphone		
McDonalds						
Öfferl (3)						
Starbucks (16)						
Vapiano (2)						
Volksbank (2)						
Antique shop						
Backery						
Card shop on corner						
Hanfshop						

Individual correlation between the aesthetic evaluations and the viewing time

Beauty:

ID	COR	P-value	Confidence interval
1	-0.2250175802	0.340173274	-0.60708764
			0.24155614
2	-0.3462963096	0.134733017	-0.68399884
			0.11363887
3	-0.5678256754	0.009010715	-0.80745378
			-0.16735762
4	0.1422949635	0.549542610	-0.32040098
			0.55017224
5	0.0032478604	0.989157419	-0.43990517
			0.44512888
6	0.1272478789	0.592920097	-0.33408518
			0.53939540
7	-0.0961211402	0.686857313	-0.51666525
			0.36178854
8	0.1574967799	0.507232736	-0.30637714
			0.56092365
9	-0.0006362627	0.997875858	-0.44303230
			0.44200897
10	-0.0863312491	0.717425196	-0.50939156
			0.37033768
11	-0.0709730067	0.766207989	-0.49785756
			0.38359537
12	-0.0325103194	0.891770482	-0.46829399
			0.41599517
13	0.3053137875	0.190532112	-0.15864055
			0.65882257
14	-0.2059079972	0.383785189	-0.59427893
			0.26033405

15	0.3940645121	0.085577618	-0.05869089
			0.71236232
16	-0.4331406961	0.056428872	-0.73481645
			0.01160433
17	-0.0541142048	0.820743437	-0.48502037
			0.39793581
18	-0.2423204684	0.303312122	-0.61851658
			0.22424670
19	-0.2058735105	0.383866604	-0.59425564
			0.26036762
20	0.0013943704	0.995344960	-0.44139877
			0.44364141
21	-0.1626264122	0.493318585	-0.56452111
			0.30159912
22	-0.0931114634	0.696207531	-0.51443557
			0.36442500

Interest:

ID	COR	P-value	Confidence interval
1	-0.182913228	0.44018064	-0.57860041
			0.28247163
2	-0.379709657	0.09867853	-0.70394669
			0.07549681
3	-0.515753521	0.01992987	-0.78020649
			-0.09488953
4	0.137066049	0.56445325	-0.32517828
			0.54644259
5	0.053728880	0.82200021	-0.39826101
			0.48472478
6	0.275333877	0.24002369	-0.19038340

			0.63988986
7	-0.018792415	0.93732112	-0.45750854
			0.42728166
8	0.093825592	0.69398505	-0.36380009
			0.51496514
9	-0.246958547	0.29386134	-0.62155332
			0.21955629
10	-0.249029123	0.28970119	-0.62290538
			0.21745536
11	-0.023736675	0.92087473	-0.46141081
			0.42322970
12	0.017958412	0.94009790	-0.42796338
			0.45684863
13	0.418535389	0.06627020	-0.02943754
			0.72650056
14	-0.213951041	0.36506250	-0.59969411
			0.25247336
15	0.526598771	0.02053848	0.09514742
			0.79149378
16	-0.357364084	0.12189391	-0.69066275
			0.10115316
17	-0.190573599	0.42092917	-0.58385617
			0.27515147
18	-0.171914446	0.46860932	-0.57099628
			0.29288800
19	0.005825418	0.98055377	-0.43782402
			0.44719339
20	-0.062406589	0.79380360	-0.49135793
			0.39090964
21	-0.093441890	0.69517890	-0.51468064
			0.36413590
22	0.045463419	0.84905467	-0.40520958

			0.47836031

Meaning:

ID	COR	P-value	Confidence interval
1	0.00000000	1.00000000	-0.44252078
			0.44252078
3	-0.12931717	0.58687063	-0.54088549
			0.33221478
5	-0.52293887	0.01798902	-0.78402694
			-0.10463085
7	0.12519975	0.59893328	-0.33593289
			0.53791799
9	-0.09844938	0.67965398	-0.51838620
			0.35974397
11	-0.20137364	0.39457237	-0.59121046
			0.26473853
13	-0.14634687	0.53811190	-0.55305121
			0.31668278
15	-0.03287735	0.89055559	-0.46858080
			0.41569128
17	-0.01901602	0.93657675	-0.45768539
			0.42709879
19	-0.20912467	0.37623392	-0.59644887
			0.25719770
21	-0.08871433	0.70994407	-0.51116772
			0.36826373
23	0.08592006	0.71871855	-0.37069507
			0.50908472
25	0.33317712	0.15115380	-0.12825300
			0.67602610

27	-0.26473292	0.25933631	-0.63308761 0.20137942
29	0.38232428	0.09618978	-0.07245489 0.70548634
31	0.03676231	0.87771116	-0.41246854 0.47161089
33	-0.10882096	0.64789631	-0.52601138 0.35058230
35	-0.17191445	0.46860932	-0.57099628 0.29288800
37	0.30103669	0.19713464	-0.16322858 0.65614864
39	0.06909354	0.77224050	-0.38520501 0.49643564
41	0.09498253	0.69038959	-0.36278681 0.51582240
43	-0.08516533	0.72109443	-0.50852125 0.37135070

Liking :

ID	COR	P-value	Confidence interval
1	-0.1508565460	0.525519545	-0.55624406 0.31252770
2	-0.4103211212	0.072343844	-0.72178360 0.03934348
3	-0.5841301589	0.006842771	-0.81578015 -0.19097426
4	0.1905203457	0.421061420	-0.27520255 0.58381974
5	0.1322217667	0.578423604	-0.32958322 0.54297274

6	0.2178976690	0.356069832	-0.24859356 0.60233842
7	0.0011398349	0.996194708	-0.44160369 0.44343695
8	0.0008998057	0.996996032	-0.44179689 0.44324410
9	-0.2004851943	0.396705413	-0.59060791 0.26559925
10	-0.1066305866	0.654556843	-0.52440659 0.35252451
11	0.2092227390	0.376005024	-0.25710193 0.59651494
12	0.0741274685	0.756112336	-0.38088755 0.50023894
13	0.1397139773	0.556880273	-0.32276200 0.54833332
14	-0.2724036918	0.245266465	-0.63801536 0.19343452
15	0.5057012178	0.022921210	0.08139536 0.77482835
16	-0.3041823962	0.192263967	-0.65811612 0.15985615
17	-0.1047090325	0.660420735	-0.52299630 0.35422510
18	-0.1719144465	0.468609323	-0.57099628 0.29288800
19	0.0040822788	0.986372038	-0.43923197 0.44579773
20	-0.0957557894	0.687990054	-0.51639489 0.36210897
21	-0.0981271796	0.680649245	-0.51814825 0.36002718

22	-0.0136854414	0.954334368	-0.45346003
			0.43144823

Attractiveness:

ID	COR	P-value	Confidence interval
1	-0.1832611443	0.43929642	-0.57883983
			0.28214033
2	-0.1221587149	0.60790777	-0.53571969
			0.33866980
3	-0.5011211472	0.02439708	-0.77236496
			-0.07529824
4	0.1816013371	0.44352317	-0.28371984
			0.57769702
5	0.0768880179	0.74730843	-0.37851144
			0.50231767
6	0.1371817021	0.56412154	-0.32507287
			0.54652526
7	-0.1342102626	0.57267106	-0.54439875
			0.32777750
8	0.0453937011	0.84928360	-0.40526796
			0.47830643
9	-0.0652940539	0.78447388	-0.49355410
			0.38845061
10	-0.0672898038	0.77804182	-0.49506885
			0.38674720
11	-0.0007018104	0.99765703	-0.44308499
			0.44195623
12	0.0518192343	0.82823482	-0.39987103
			0.48325838
13	0.4895643550	0.02845651	0.06005376

			0.76611265
14	-0.2587151930	0.27072425	-0.62920076
			0.20756958
15	0.4607153659	0.04092030	0.02285396
			0.75027333
15	-0.2667409651	0.25560482	-0.63438048
			0.19930553
16	-0.0607663236	0.79911560	-0.49010793
			0.39230365
17	-0.1085694606	0.64865977	-0.52582727
			0.35080551
18	-0.1017588818	0.66946064	-0.52082662
			0.35683011
19	-0.0355912918	0.88157983	-0.47069862
			0.41344114
20	0.0362947967	0.87925536	-0.41285696
			0.47124679
21	0.0134721836	0.95504527	-0.43162181
			0.45329057

Emotional arousal:

ID	COR	P-value	Confidence interval
1	-0.2506938902	0.286382795	-0.62399084
			0.21576305
2	-0.2744307935	0.241631755	-0.63931260
			0.19132472
3	-0.4543476328	0.044167525	-0.74673181
			-0.01480315
4	0.0864609480	0.717017391	-0.37022492
			0.50948832
5	0.2221813897	0.346455059	-0.24436537

			0.60519907
6	0.3350021308	0.148789953	-0.12623196
			0.67714002
7	-0.3362216358	0.147224830	-0.67788348
			0.12487932
8	0.0688498977	0.773023457	-0.38541347
			0.49625114
9	-0.0005150084	0.998280661	-0.44293484
			0.44210653
10	-0.1119139858	0.638534752	-0.52827246
			0.34783297
11	0.0142876548	0.952327041	-0.43095789
			0.45393837
12	-0.1138707454	0.632639262	-0.52969986
			0.34608955
13	0.4055386172	0.076066480	-0.04507048
			0.71902390
14	-0.3789422826	0.109604799	-0.71080794
			0.09091526
15	0.5730738743	0.008259214	0.17490968
			0.81014439
16	-0.1614239482	0.496563411	-0.56367918
			0.30272126
17	-0.2133009179	0.366556169	-0.59925771
			0.25311104
18	-0.1094802353	0.645896568	-0.52649382
			0.34999695
19	0.0720864637	0.762640206	-0.38264046
			0.49869887
20	0.0887005549	0.709987224	-0.36827573
			0.51115747
21	-0.0187491705	0.937465081	-0.45747433

			0.42731702
22	0.0507198732	0.831828534	-0.40079663
			0.48241309

Familiarity:

ID	COR	P-value	Confidence interval
1	0.208382467	0.37796879	-0.25792225
			0.59594869
2	0.106611266	0.65461571	-0.35254162
			0.52439242
3	-0.401305315	0.07947905	-0.71657284
			0.05011523
4	0.148805010	0.53123092	-0.31442012
			0.55479305
5	0.009039095	0.96982982	-0.43522257
			0.44976084
6	0.185070162	0.43471389	-0.28041596
			0.58008361
7	-0.310585468	0.18260106	-0.66210605
			0.15295793
8	0.195841350	0.40795763	-0.27008614
			0.58745133
9	0.026692720	0.91105621	-0.42079857
			0.46373582
10	-0.352862744	0.12700564	-0.68795930
			0.10624867
11	-0.111403969	0.64007483	-0.52790003
			0.34828686
12	0.129458471	0.58645850	-0.33208693
			0.54098715
13	0.227663218	0.33437411	-0.23892859

			0.60884544
14	0.043995920	0.85387600	-0.40643785
			0.47722555
15	0.169602251	0.47470130	-0.29506381
			0.56938896
16	-0.287396106	0.21920592	-0.64756090
			0.17772789
17	-0.084927972	0.72184213	-0.50834397
			0.37155680
18	0.106676026	0.65441841	-0.35248425
			0.52443991
19	0.292522860	0.21072845	-0.17230197
			0.65079933
20	0.121084194	0.61109187	-0.33963500
			0.53494159
21	0.252452106	0.28290371	-0.21397270
			0.62513567
22	0.116268384	0.62544446	-0.34394898
			0.53144566

Zusammenfassung

Das natürliche Sehverhalten wird meist im Labor untersucht, wo Bilder oder Videos für einige Sekunden auf einem Computerbildschirm erscheinen. Doch was sagt dies über das Blickverhalten unter realen Bedingungen aus? Diese Arbeit untersucht, wie Menschen im Alltag mit Text umgehen. Untersuchungen im Labor konnten bereits zeigen, dass Text generell Aufmerksamkeit erregt, indem sie Bilder von realen Szenen verwendeten (Cerf, Frady & Koch, 2009). Das umgesetzte Studiendesign geht aus einem freien Erkundungsparadigma am Donaukanal in Wien mit urbaner Kunst heraus (Mitschke, Goller & Leder, 2017). In dieser Studie wurde ein Mixed-Method-Design verwendet, das Mobile-Eye-Tracking und Fragebögen kombinierte. Zunächst wurde eine mobile Eye-Tracking-Feldstudie durchgeführt. In einem Follow-Up Fragebogen wurden die Teilnehmer nach der Erinnerung, dem Wiedererkennen und der ästhetischen Bewertung von Textelementen, die ihnen während des Spaziergangs begegneten, befragt. Die Ergebnisse der Mobile-Eye-Tracking-Analyse zeigten, dass Text mit 24.37% der durchschnittlichen Fixierungsdauer die größte Aufmerksamkeit auf sich zog, gefolgt von Personen und Natur. Diese Ergebnisse sind signifikant, da die Teilnehmer nicht angewiesen wurden, sich auf Text zu konzentrieren.