Perceived environments in reality and online

Name: Tomáš Pudil

Student number: 0809756
Master programme: Toxicology and Environmental Health
Institute: Utrecht University, Faculty of Veterinary Medicine, the Institute for Risk Assessment Sciences (IRAS), One Health Epidemiology Chemical Agents

Context: This project is the author's major research project as part of the master programme Toxicology and Environmental Health.

Daily supervisor: K. Kyriakou, PhD.
First examiner: A. Huss, PhD.
Second examiner: G. Hoek, PhD.

Abstract

People's mental state can be affected by their perception of their surroundings, either propagating in other health outcomes or enhancing well-being. This perspective has driven a new discussion about the importance of research on perceived environments, particularly green spaces. The standard method in this field is the use of online surveys, gathering information on people's perceptions based on Google Street View or the Flicker database. This approach, however, does not question if these assessments limited by the online environment are an accurate representation of reality, with some elements of the actual environment not being carried over. Here, we address whether people's perceptions differ when comparing their answers from an online survey to answering the same questions in a real-life setting.

The study consisted of three typologically different locations in the city of Utrecht, the Netherlands. In total, 22 participants completed the online survey and subsequently did the assessment in the real-life setting under a researcher's supervision. The survey was based on 360-degree pictures of these locations with 20 variables to be compared. By recruiting another group of 170 and 228 participants for the online survey, it was possible to examine how different people perceive the same pictures. The pictures were analysed using a deep learning algorithm for object recognition. These data were compared to the participants' perceptions of natural-like and urban environments assessed in the real-life setting and online, resulting in calculating percentage agreements.

Six variables (Social Connection, Entertainment, Uniqueness, Colours, Cared For, Water) showed a significant ($p<0.05$) difference between the two settings. The percentage agreements exceeded consistently 50% and were as high as 88% for 'urban elements' online. Several perceptual variables developed links with the most prevalent feelings, i.e. a clustering effect of these variables, although not all held against the data from the real-life setting. This pilot study was the first attempt to examine the differences between the real-life setting and online surveys. Discovered differences do not reject using online surveys to study perceived environments, as they were less critical in the overall result interpretation. The main limitations seemed to be the temporal discrepancy of the locations captured by a camera and the assessment in the real-life setting, and the small size of the pictures in the survey layout combined with the lack of other stimuli. These limitations may be resolved by design improvements suggested in this study, yet more robust experiments are needed.

Layman's summary

Good mental health is an essential element of one's life. People's emotional states are the product of the inner self but also the interactions with the outside world. Our surroundings and the environment we live in affect us, too. That is why researchers have recently focused on the links between mood and how people perceive both urban and natural environments. Researchers tend to use online surveys for this purpose.

This study compared how people perceive different environments through pictures online to their answers to the same questions at the actual locations in a real-life setting. Of 20 questions, six of them differed in this comparison, namely the perception of social connection, entertainment, uniqueness, water-dominated environment, the pleasantness of colours, and the feeling of being cared for. Overall, the results showed minor differences, and the bulk of the measured variables remained unchanged. In some cases, the interpretation of the results was not clear, so we suggested specific improvements in the study design. Additionally, we used an algorithm based on an AI architecture to analyse our pictures. This investigation told us that the participants perceived vegetation and the urban landscape similar to the objective data provided by the picture analysis. These findings are very promising and call for further use of this technique in the future, as they bring a new layer of insights into people's subjective

perceptions. For instance, more characteristics of the environments can be included in the picture analysis with a thoughtful study design. Our study favours using online surveys in this field, although many questions still must be answered.

Content

ABSTRACT **2**

LAYMAN'S SUMMARY **2**

1 INTRODUCTION **6**

1.1 AIM 6
1.2 ETHICAL ASPECTS 6

2 MATERIAL AND METHODS **7**

2.1 TARGET POPULATION 7
2.2 CREATING THE ONLINE SURVEY 7
2.2.1 Selecting areas and locations 7
2.2.2 Capturing the locations 7
2.2.3 Picture selection 7
2.2.4 Creating an interface for participants entering the online survey and doing the real-life experience 8
2.2.5 Participant recruitment 8
2.2.5.1 Tapreaserch and vouchers 8
2.3 DATA PROCESSING 9
2.3.1 Pictures analysed by the algorithm 9
2.3.2 Data cleaning 9
2.4 DATA ANALYSIS 10
2.4.1 Reality vs Online 10
2.4.2 Correlation structures within each data set 11
2.4.3 Participants' perception vs image mining algorithm 11
2.4.4 Reliability of the paid campaign 11

3 RESULTS **13**

3.1 POPULATION DESCRIPTION 13
3.2 DESCRIPTIVE STATISTICS 13
3.3 REALITY VS ONLINE 13
3.3.1 Net values 13
3.4 CORRELATION STRUCTURES WITHIN EACH DATA SET 15
3.5 PARTICIPANTS' PERCEPTION VS IMAGE MINING ALGORITHM 16
3.6 RELIABILITY OF THE PAID CAMPAIGN 17
3.6.1 Cen online vs Cen campaign 17
3.6.2 Cen campaign and Ove campaign 17

4 DISCUSSION **18**

4.1 STRENGTHS 18
4.1.1 Algorithm's output 19
4.2 LIMITATIONS 19
4.2.1 Result interpretation 19
4.3 OTHER STUDIES 20
4.3.1 People's feelings towards greenery 20
4.3.2 Lower reliability for social aspects of the environment from pictures 20
4.3.3 Auditory sense 21
4.4 THE BIG PICTURE 21
4.4.1 The case against the real-life assessment 21
4.4.2 Subjective experience vs objective properties of a place 22
4.5 OPPORTUNITIES 22

5 CONCLUSION **23**

6 APPENDICES **24**

6.1 APPENDIX 1 – DEMOGRAPHICS AND DESCRIPTIVES 24

6.2 Appendix 2 – Reality vs Online ... 26
6.2.1 Results of the Wilcoxon test ... 26
6.2.2 Differences between the Cen locations within three data sets ... 26
6.3 Appendix 3 – Correlation structures ... 28
6.4 Appendix 4 – Reliability of the paid campaign ... 32
6.5 Appendix 5 – Image mining algorithm ... 33
6.6 Appendix 6 – Search for new locations ... 34

REFERENCES ... 35

1 Introduction

Apart from pollutants and toxins in the environment, another area of interest in exposure science is to study green spaces that also contribute to people's health and well-being. One of the first attempts regarding this topic dates back to a classic study by Ulrich, 1984, on a view through the hospital window (1). More recent findings have shown multiple health benefits associated with exposure to green (2–4). Greenery is, without a doubt, part of perceived environments and the urban landscape (5,6). Most humans encounter it daily. Current approaches to study perceived environments vary among disciplines, from geolocated questionnaires capturing real-time perceptions and subjective emotions to incorporating image mining and virtual reality (7–10). A new technological development has led to objective study setups since wearable devices measuring physiological signals can be used (10,11). This development embarked on new opportunities to track and ask people in-depth questions about their feelings or perceptions, whether they are on a busy street or trying to escape the hectic pace of city life in natural environments.

There has been a tendency to use online surveys based on pictures of the environment or street views with the ultimate goal of quantifying big data sets and predicting visual cues or emotions specific to certain environments (8,12). This approach can be problematic because the pictures only display fragments of reality which could, in turn, lead to wrong assumptions about the environment. For instance, environmental conditions, e.g. weather, crowdedness, vegetation cover, might change so that the same scene will demonstrate different qualities, potentially affecting people in a new way. Under this assumption, pictures in the online surveys may not fully represent the real-life setting due to the ever-changing environmental factors and the lack of other stimuli such as sound and smell (13–15). Thus, it is worth exploring the discrepancies between the online surveys and real-life setting. Another question regards the differences among people assessing the same pictures and what factors may affect those variations.

1.1 Aim

The primary aim of this project is to examine whether there are differences in participants' perceptions and feelings through pictures of various locations in an online survey compared to the same questions subsequently done in a real-life setting.

The secondary aim is to identify if different people will rate the same pictures similarly or not and look for explanatory factors of these deviations.

1.2 Ethical aspects

Participants in the online survey were informed on the ethical nature of this study, with a disclosure that all sensitive personal information will be stored securely. Participants were mostly volunteers, and those doing the assessment in the real-life setting were given financial compensation after signing consent.

2 Material and Methods

The project consisted of two main phases: 1) creating an online survey 2) doing an assessment in a real-life setting with the same participants thereafter. The online survey was divided into separate versions for each area, asking the same questions but gathering data from different groups of people. Along with the project, different strategies were used to invite participants.

2.1 Target population

The surveys targeted adults living in the Netherlands. No other characteristic was required.

2.2 Creating the online survey

The survey was created on the platform Qualtrics. The basis of the survey came from a questionnaire used in a previously conducted research, where a customised subjective perception questionnaire (SPQ) was created based on the existing literature as one of the metrics for studying stress levels in humans in urban environments (16,17). The survey was split into three main sections: demographics, perception of the conditions of a given picture, and feelings. The questionnaire was adapted to the needs of the current study. One additional question was added, specifically assessing the overall physical appearance of a place defined by three environmental types: a natural-like, urban, water-dominated environment. Answer options of some questions were changed from using a scale to selecting multiple answers. The opinion section used a bipolar scale, from -2 as very unpleasant to +2 as very pleasant. The feelings were on the unipolar scale from nothing felt (0) to very intense (4). The data collection took almost three months, starting on May 19th and ending on August 10th, 2021.

2.2.1 Selecting areas and locations

The previous study of stress induction in urban environments consisted of six routes of six stops (all in the city of Utrecht), making up for 36 locations. Out of those, 9 locations were selected, resulting in three areas, each consisting of three locations. Even with an initial attempt to add more locations to the current collection, the selection for the survey was narrowed, represented by two residential areas outside the city centre and one in the historic city centre. The main assessment comparing the online survey and real-life setting was conducted only for one area. The researcher's selection of these locations was based on walkability within one area, the representation of different environments (natural-like with a lot of vegetation, semi-urban to urban, and a water-dominated environment), the distance from and similarities with the other locations.

2.2.2 Capturing the locations

A camera, GoPro Max, was used to capture our locations, having a collection of 360-degree images taken during the previous study. The scenes were captured under the same procedure, sometimes adjusted due to changed conditions. The entire picture collection took place from October 14th to December 16th, 2020. In the current project, new pictures of three locations used in the survey were re-taken at the end of the summer of 2021. The pictures were analysed with an algorithm for object recognition (18), provided by the university's IT department.

2.2.3 Picture selection

The pictures used in the online survey were taken during the previous project. Pictures taken in dark, dim conditions or overexposed were excluded. To summarise, for the total of nine locations, pictures used in the online survey were taken on eight different days during the autumn and winter of 2020.

2.2.4 Creating an interface for participants entering the online survey and doing the real-life experience

An interface for entering the online survey was created on Google maps to share it efficiently. This approach had the advantage of using only a single link for all the areas. Participants could choose from the three areas, intentionally covering a larger space to reduce the bias about specific neighbourhoods. Lastly, participants were given an option to continue assessing more locations online when they opted not to do the assessment in the real-life setting thereafter.

Fig. 1: The entry interface on Google maps; A) Three areas (LeR – blue on the left, Cen – blue in the centre, Ove – red) from the online surveys, B) A detailed map view of exact locations exemplary shown for the Cen area.

Fig. 2: Pictures of the three locations from the Cen area used in two online surveys (Cen online, Cen campaign) and assessed in the real-life setting; A) CeT3 – a location representing an urban environment, B) CeT6 – a city park with some urban elements, C) CeB5 – a city park with the most amount of vegetation.

2.2.5 Participant recruitment

Word of mouth was also used as recruiting strategy. A link for the map (the interface for entering the survey) with a short welcoming message was used to share the survey on social media (posting in private chats or student groups on WhatsApp). A poster was also shared through the Veterinary Faculty newsletter and its social media, and a Reddit group called "Not Just Bikes", devoted to a YouTube channel of the same name about urbanism in the Netherlands.

The initial idea for marketing the survey was first to attract people to participate in the online survey with an opt-in option to do the assessment in a real-life setting. A single survey link for the map view (the interface for entering the survey) with a short welcoming message was used for most sharing purposes of the survey.

2.2.5.1 Tapreaserch and vouchers

A modified version of the online survey was used in a paid campaign to have enough participants. Tapresearch was selected for two reasons: 1) their novel strategies to target real responders 2) the service geolocated for The Netherlands. Tapresearch uses a mobile

environment to attract responders through a game or an app they are using. The anonymous link from Qualtrics was edited to track the target number of campaign participants. Each of the three areas had a separate campaign, which made up for expenses of just over 150 euros.
The survey was also marketed on Proefbunny, a website for sharing research dedicated to psychology, medical- neuroscience, and recruitment of paid participants. After completing the online survey and the real-life assessment, the participants were compensated with a 5 euro voucher. The same applied to participants recruited by other means.

2.3 Data processing

2.3.1 Pictures analysed by the algorithm

A deep learning algorithm analysed the pictures to assign each pixel with one of 19 recognisable objects. The objects were divided into two categories, 'vegetation' and 'urban elements', corresponding to the variables 'Green' and 'Grey' in the survey that represented natural-like and urban environments, respectively. The vegetation category contained two objects: vegetation and terrain; the remaining sections, apart from 'sky', fell under 'urban elements'. In the final output, the user was presented with percentages of every object. More information about the picture analysis process can be found in the discussion section under Algorithm's output and related Fig. 14 (Appendix 5).

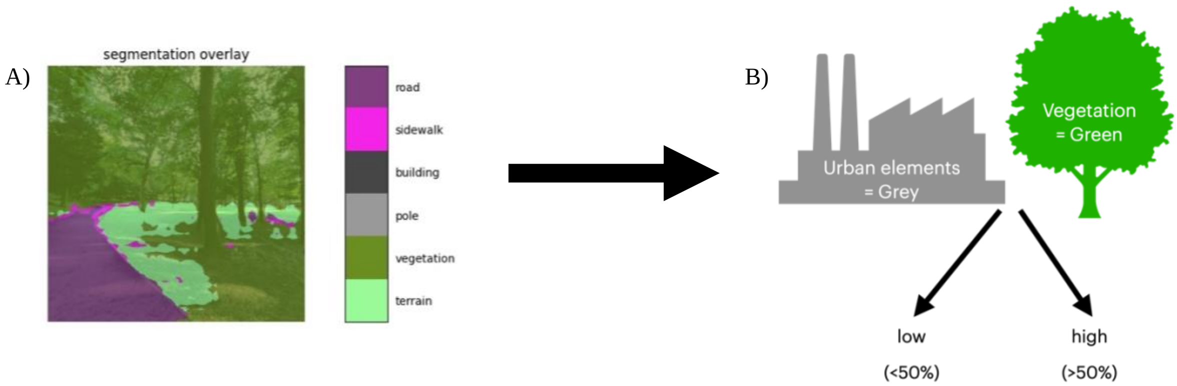

Fig. 3: A process diagram of the object recognition analysis; A) The overlay only shows some of the potentially recognised objects, B) Two categories named 'vegetation' and 'urban elements' representing the variables 'Green' and 'Grey' further divided into two categories based on the percentages calculated by the algorithm.

The following are all potentially recognised objects: bicycle, building, bus, car, fence, motorcycle, person, pole, rider, road, sidewalk, sky, terrain, traffic light, traffic sign, train, truck, vegetation, and wall.

2.3.2 Data cleaning

Original data came from multiple versions of the survey. Overall, we gathered data from six online and nine real-life versions of the survey. *Cen online* (the city centre area) was the only one with a sufficient number of participants (N=22) who did both the online and the real-life assessment, and hence was selected for further analyses.
The same steps were followed for three data sets of the online paid campaign - Cen (N=279), Ove (N=372), LeR (N=144). Before downloading, they were prefiltered for computer bot recognition in Qualtrics, counting for three findings in *Cen campaign* and one in the Ove and LeR data sets. Of them, Cen (from now on *Cen campaign*) and Ove (from now on *Ove campaign*) were selected: 1) *Cen campaign* for the ability to be compared with *Cen online* where the same pictures were assessed, and *Cen reality* representing the actual locations in the real-life setting, 2) *Ove campaign* for its size. A further data cleaning of these two data sets was done in R using the following criteria: the duration (>4 min), otherwise answering the

questions would not be feasible (tested by the researcher), the requirement to live in NL, the completion (>90%), and lastly looking for straight liners by summing up all numeric variables and excluding participants that completed the survey with the highest and lowest ratings. After the cleaning, *Cen campaign* consisted of 170 and *Ove campaign* of 228 participants.

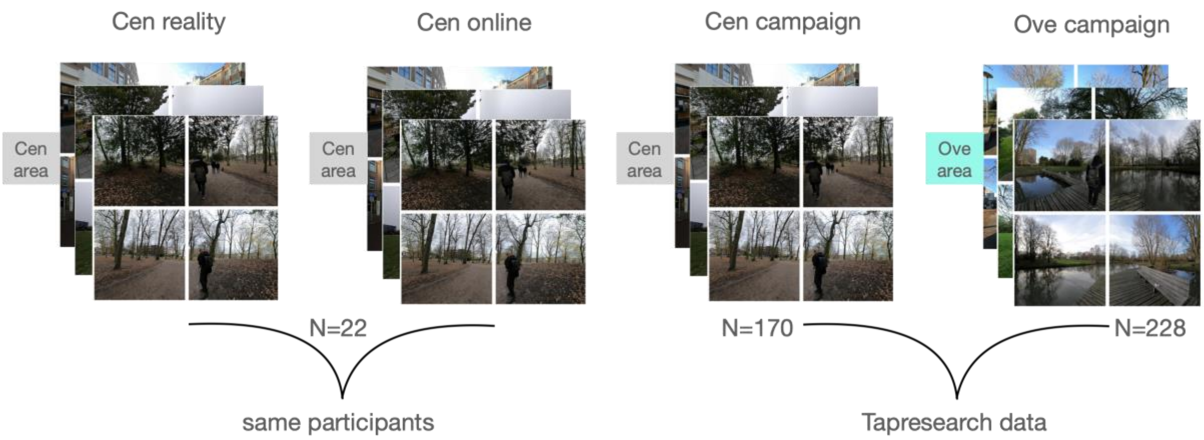

Fig. 4: An overview of the data processing of four analysed surveys, distinguishing different recruitment strategies and areas assessed in Utrecht online or in the real-life setting.

2.4 Data analysis

Demographic data were calculated for three groups of participants, then multiple tables containing min, max, median, mean, sd, 25^{th} and 75^{th} percentiles, and the IQR were processed for every location and setting separately. It is worth noting that the real-life setting (*Cen reality*) was used as the gold standard and is mentioned in all tables and graphs first, even though the participants first filled out the online survey (*Cen online*). All steps of the statistical analyses are listed in Tab. 1, with the major ones summarised in a flow chart (Fig. 5) further below.

Tab. 1: Different statistical methods applied in this study and the purpose of their use; statistical significance was set at p<0.05 for all tests.

Purpose	Method
Normality of data	The Shapiro-Wilk test
Reality vs Online	The Wilcoxon sum rank test, The Spearman correlation
Correlation structures	The Spearman correlation
Difference between the locations	The Kruskal-Wallis test with a post-hoc analysis
Participants' perception vs image mining algorithm	The Percentage agreement, The Kappa coefficient
Reliability of the paid campaign	The Wilcoxon sum rank test, The Kruskal-Wallis test

2.4.1 Reality vs Online

Normality of the data was tested by the Shapiro-Wilk test and using plots such as 'ggdensity' and 'qqplot' for all numeric variables to decide which statistical tests to use. None of the tested variables (N=20, all numeric variables in the survey) showed normal distribution, and therefore further analyses were performed using non-parametric statistical tests.

The next step was to compare those 20 variables between *Cen reality* and *Cen online* (N=66) per the whole data set, i.e. all locations combined, using the Wilcoxon rank-sum test with continuity correction, the unpaired version. The Spearman correlation was used to look for correlations between the same variables in the real-life setting and online. The correlations were processed for all numeric variables of *Cen reality* and *Cen online* per location (N=22), for example, to see how the opinion on maintenance in the real-life setting correlated with the participants' perception of this variable online. All significant correlation estimates were considered relevant for further result interpretation.

The Kruskal–Wallis test was performed with the variable 'Location' as a grouping factor to look for differences of the same variables across the locations in the *Cen area*. This analysis was processed within each data set consisting of the same pictures, as in the case of *Cen online* and *Cen campaign*, and *Cen reality* which represented the real-life setting of these pictures. Then a pos-hoc analysis was performed for the variables with the highest chi-squared values that appeared the most frequently.

2.4.2 Correlation structures within each data set

The Spearman correlation was used again to look for relationships between the feelings and perceptions developed in the real-life setting and online within each data set, i.e. all locations combined. In total, four data sets were used in this analysis: *Cen reality* (N=66, the real-life setting), *Cen online* (N=66, the same participants online), *Cen campaign* (N=510, the same pictures rated by different people online), *Ove campaign* (N=684, different locations rated by another group of people online).

2.4.3 Participants' perception vs image mining algorithm

A percentage agreement was performed to compare the participants' perception of a natural-like and urban environment (the variables 'Green' and 'Grey') with the algorithm's output. For this purpose, the answers from the surveys were turned into two categorical values (from the original scale of 0-4) – 'low' (up to 2 points) and 'high' (3,4) in comparison with 'vegetation' and 'urban elements' recognised by the algorithm, both with two categories of 'low' (<50%) and 'high' (>50%). The last step before the actual calculation was to assign the two categories with a numeric value of '0' for 'low' and '1' for 'high'. The Kappa coefficient was processed to verify the percentage agreements excluding *Ove campaign* for its invariability.

2.4.4 Reliability of the paid campaign

Differences between the two online data sets assessing the same locations (*Cen online*, *Cen campaign*) were tested statistically by running the Wilcoxon rank-sum test per data set, i.e. all locations combined. That way, the data from the paid campaign (*Cen campaign*) were compared to the participants recruited by normal means (*Cen online*).

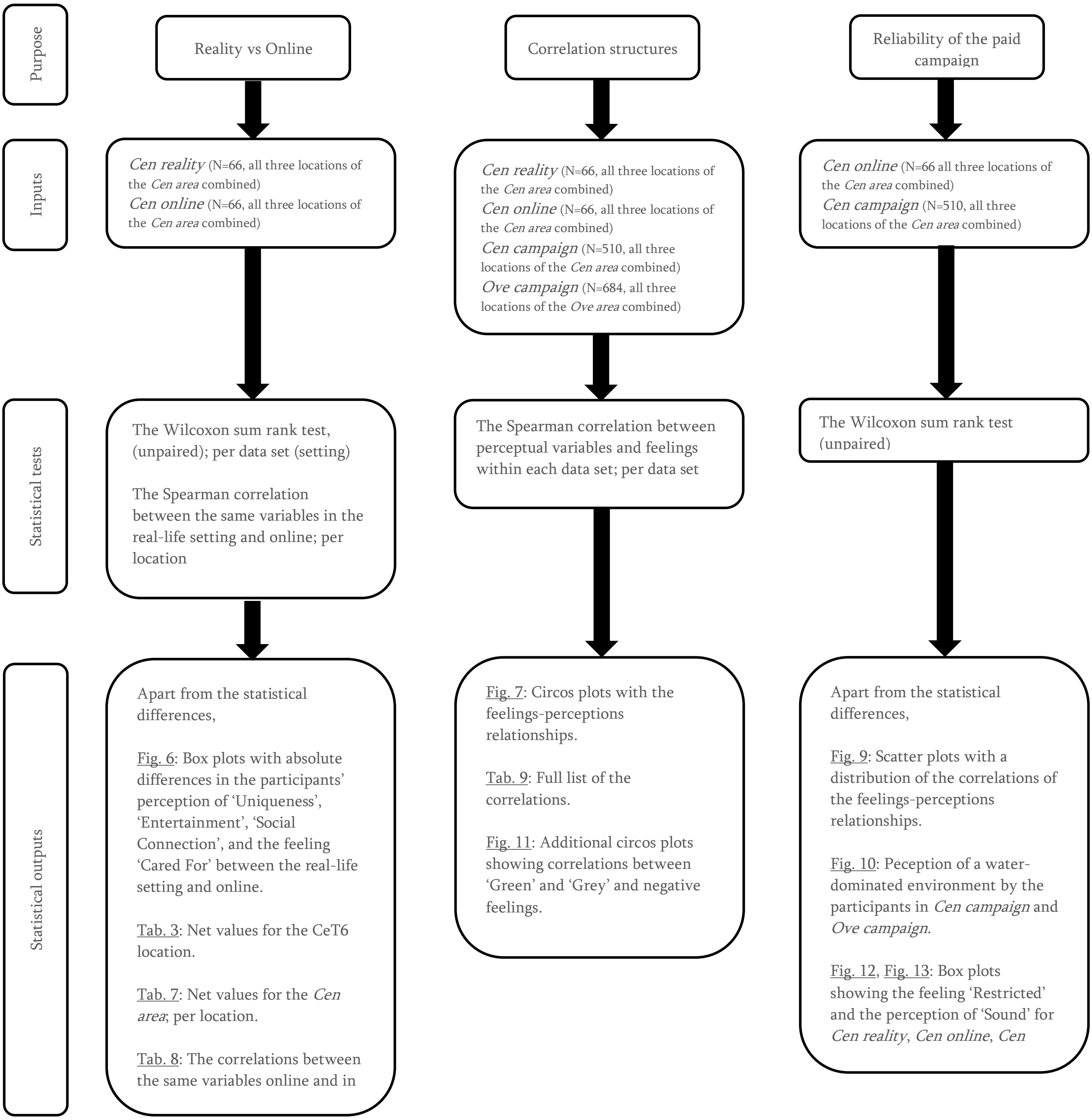

Fig. 5: A Flow chart summarising the most important parts of the data analysis section, listed with the purpose of the procedure, input data, statistical methods used, and statistical outputs. Note that the comparison of the participants' perception and the image mining algorithm is discussed separately.

3 Results

3.1 Population description

The online data collection counted for 45 participants in Cen, 22 in Ove, and 34 in LeR. The *Cen area* was the only one with a sufficient number of participants for the real-life setting. The group of 22 participants was represented mainly by males (N=17), young adults with higher education living in Utrecht. Compared to the two data sets of the paid campaign, some differences can be found in every category. In contrast with the primary data set, both *Cen campaign* and *Ove campaign* differed in the ratio of university graduates, the place of participants' residency and their upbringing. The complete demographic data with additional text can be found in Appendix 1.

Tab. 2: Demographic information on the main participants (N=22) who did both the online survey and the assessment in the real-life setting in the Cen area.

Groups	Categories	Count	Prop
Gender	Male	17	77.27
	Female	5	22.73
Age	18 - 30	16	72.73
	31 - 50	4	18.18
	51 - 65	2	9.09
Education	High School	1	4.55
	MBO	2	9.09
	University	19	86.36
Childhood in NL	Yes	8	36.36
	No	14	63.64
Living in Utrecht	Currently living	17	77.27
	Yes, in the past	3	13.64
	No, never	2	9.09

3.2 Descriptive statistics

Regarding the data from the participants who did the assessment in the real-life setting and online survey, the three most pronounced feelings in both settings (*Cen reality* and *Cen online*) were 'Safe', 'Calm', 'Free'. Negative feelings consistently scored the lowest where the common ones were 'Rejected', 'Threatened', 'Anxious', 'Irritated', with their mean values ranging from 0.18 to 0.68 for both settings. Besides 'Uniqueness' and 'Entertainment', which showed negative mean values for certain locations online, the perceptual variables had lesser importance in this section. Tables with all descriptives and additional text can be found in Appendix 1. *Cen campaign* and *Ove campaign* followed the same patterns regarding the mean values of positive and negative feelings and perceptions.

3.3 Reality vs Online

Statistical tests comparing the two settings resulted in only a few differences. Notably, they included one feeling ('Cared For'), followed by five other perceptual variables that were also significantly different ($p<0.05$). The participants perceived the locations as more unique, entertaining, and social in the real-life setting, while the feeling of being cared for implied the opposite.

Regarding the Spearman correlations, the significantly correlated variables for each location were 'Threatened', 'Rejected', and 'Water'. On the other hand, the five following variables 'Social Connection', 'Colours', 'Irritated', 'Stimulated', and 'Free' did not show any significant correlation. No negative correlation was found with a significant p-value.

Tables and the full description of the statistical differences and correlations are listed in Appendix 2.

3.3.1 Net values

On average, 'Social Connection', 'Uniqueness', and 'Entertainment' have a net positive score in the real-life setting across all the locations, illustrated with Fig. 6, meaning that the participants were more inclined to give higher ratings of these variables in the real-life setting. In terms of the differences in the rating between the settings, Tab. 7 in Appendix 2 shows net mean values for each location separately, primarily pointing to negative feelings whose means were, in most cases, close to 0, reflecting a very slight difference between the answers in the real-life setting and online. Overall, 'Uniqueness' had the highest mean of 0.6, whereas 'Cared For' scored the lowest with the mean of -0.6. Notably, all negative feelings at the CeT6 location were rated higher in the real-life setting than online, and the opposite applied to most positive feelings showing that this location induced contrary feeling tones in the two settings. Moreover, CeT3 was perceived as more entertaining (1.0) but also less calm (-0.5) in the real-life setting, and at the last location, CeB5, the participants felt safer (1.0), but at the same time, less cared for (-0.9) than online.

Tab. 3: Differences between the ratings in the real-life setting and online for the CeT6 location, where all negative feelings gained higher ratings in the reality.

	CeT6 location		
Variables	**mean**	**sd.**	**valid N**
Irritated	0.7	1.3	22
Safe	0.2	1.1	22
Restricted	0.1	1.2	22
Threatened	0.4	1.0	22
Stimulated	0.0	1.3	22
Cared For	-0.9	1.5	22
Rejected	0.3	0.6	22
Calm	-0.9	1.4	22
Free	-0.5	1.5	22
Anxious	0.1	0.8	22

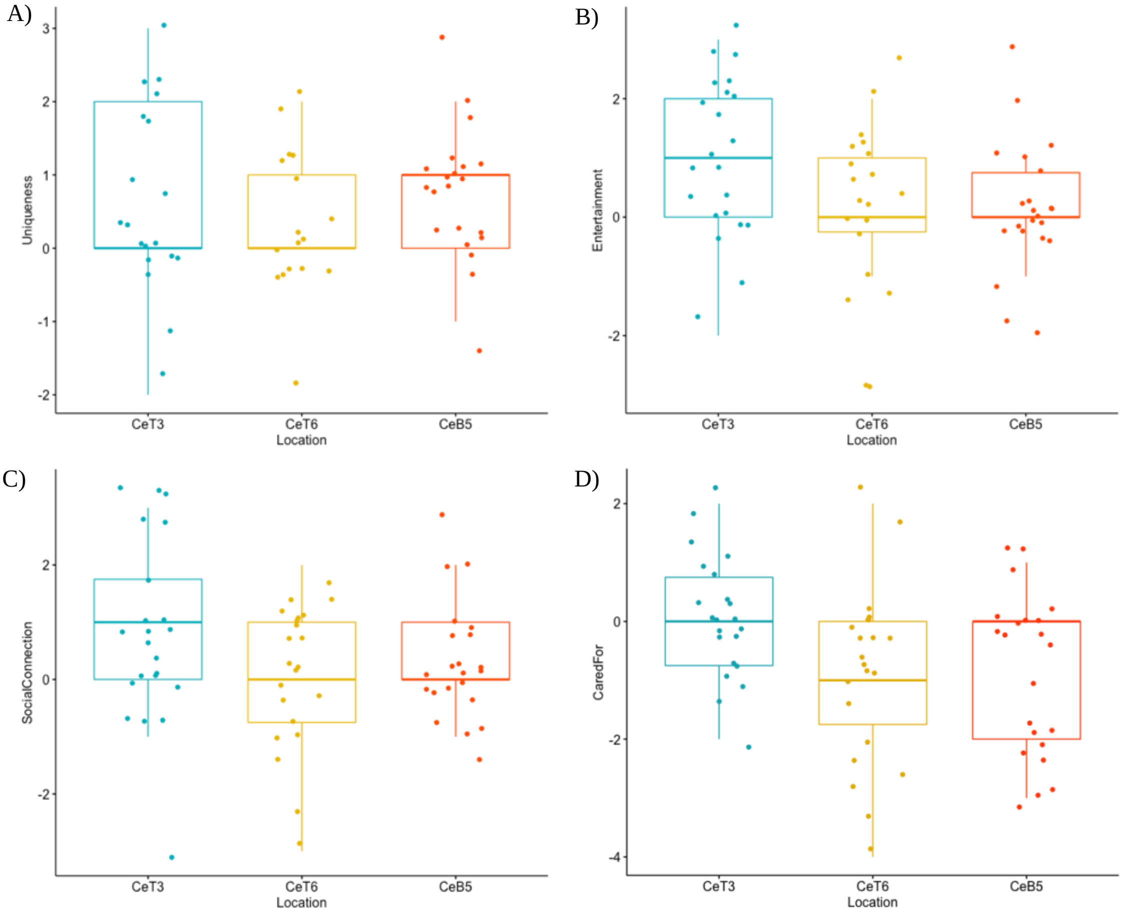

Fig. 6: The box plots show absolute differences between the ratings in the real-life setting and online for A) Uniqueness, B) Entertainment, C) Social Connection, D) Cared For (the only feeling of the four); each colour represents one location in the Cen area.

3.4 Correlation structures within each data set

Other correlations showed that some feelings, especially ‘Stimulated’, ‘Calm’, ‘Free’, were connected to multiple perceptual attributes in the real-life setting and online. Again, the feeling ‘Cared For’, which imposed the same pattern online, did not hold this clustering effect in the real-life setting. The perception of maintenance, organisation and social connection showed less clear and fewer relationships with the prevalent feelings. These feelings were highly correlated with the perceptions of ‘Uniqueness’, ‘Colours’, ‘Sound’, and ‘Green’ in all data sets. In general, the relationships indicated higher correlation values for *Cen reality* and *Cen online*, while lower ones for those from the paid campaign (*Cen campaign* and *Ove campaign*). In the real-life setting, the highest correlation was the ‘Sound-Calm’ relationship with a value of 0.59, and the overall highest one was the ‘Colours-Stimulated’ relationship with a value of 0.66 in *Cen online*. Tab. 9 in Appendix 3 shows all correlations.

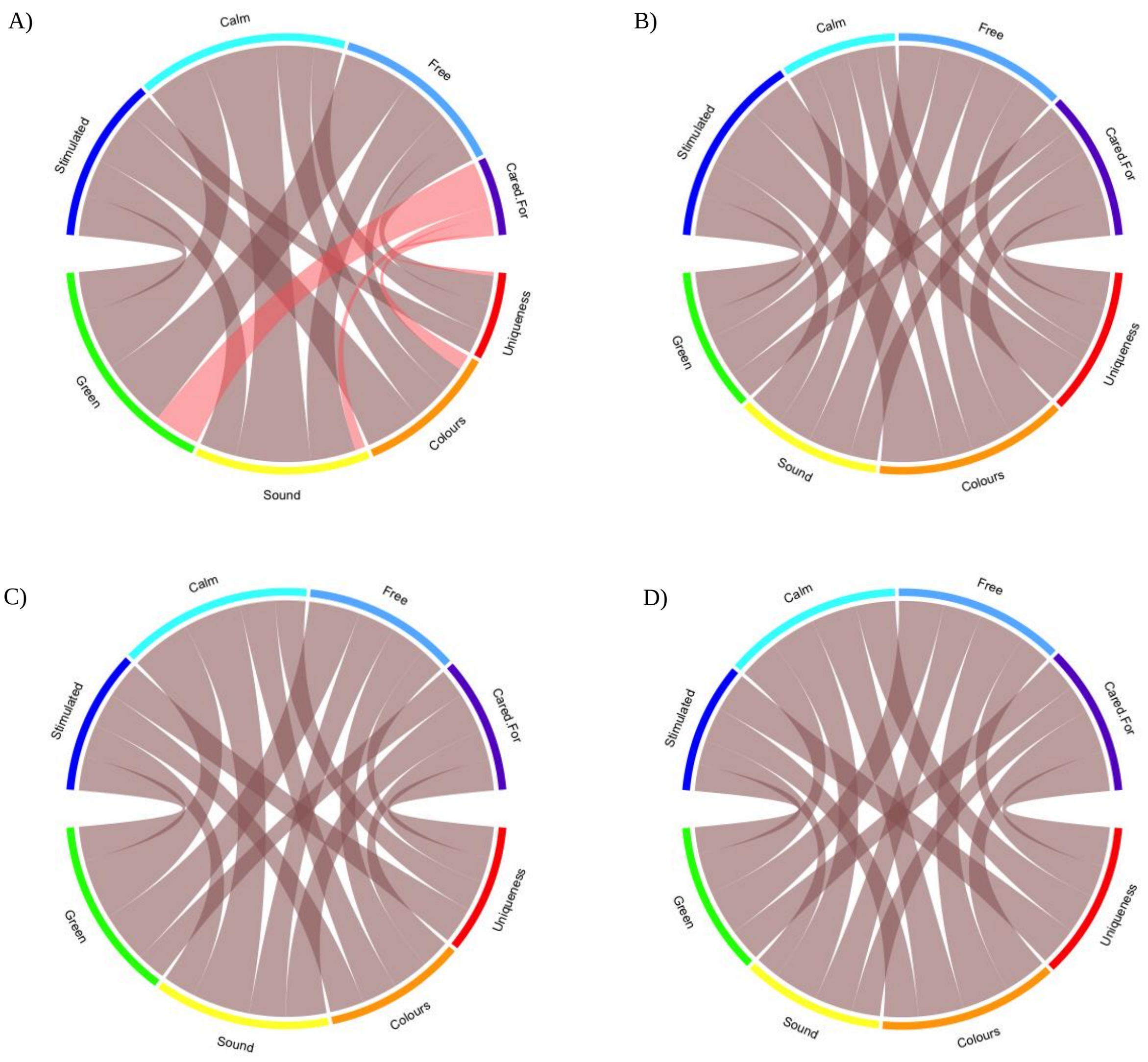

Fig. 7: Clustering effect – the relationships of multiple perceptual variables with five feelings in four surveys; A) Cen reality, B) Cen online (the same participants online), C) Cen campaign (the same pictures rated by different people online), D) Ove campaign (different pictures rated by another group of people online); the width of the links indicates its strength, i.e. the correlation coefficient; negative correlations are in red.

In Fig. 7, the perception of a natural-like environment ('Green') and the pleasantness of sound ('Sound') dominate the perceptual variables in the real-life setting, whereas online, the distribution is more even. For instance, the perception of the pleasantness of colours overtakes the key role of the relationships in *Cen online* and *Ove campaign*, wherein it has the highest correlations with the feelings. Lastly, the feeling of being cared for is much less pronounced in the real-life setting than the correlation distribution online, where this feeling always gained positive correlations as opposed to the links seen in *Cen reality*, reflecting the previously described results. The same observation was made for 'Anxious'.
It was also possible to find these correlation structures for several negative feelings with a connection to the participants' perception of a natural-like ('Green') and urban environment ('Grey'), illustrated by other circos plots (Fig. 11) with additional text in Appendix 3.

3.5 Participants' perception vs image mining algorithm

The perceptions of a natural-like ('Green') and urban environment ('Grey') were similar among the participants of the three surveys assessing the same locations (*Cen reality*, *Cen online*, *Cen campaign*). The box plots below (Fig. 8) illustrate these findings, where the only noticeable difference was in the perception of 'Green' for the last two locations in *Cen campaign* compared to the assessment done by the main participants (*Cen reality, Cen online*).

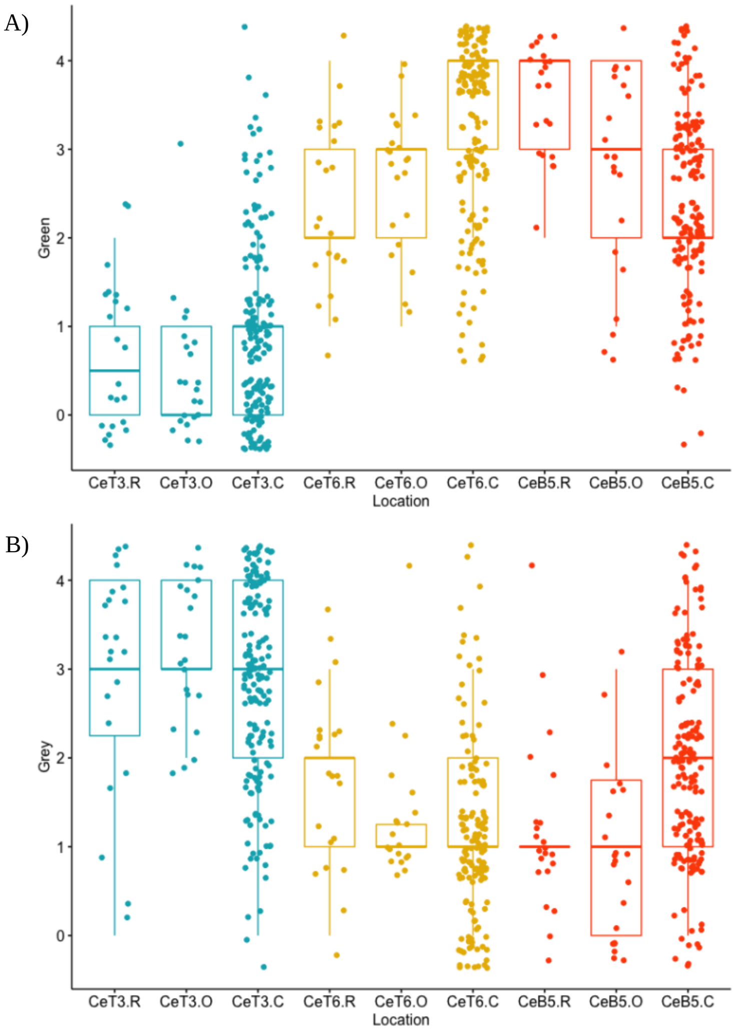

Fig. 8: The two box plots show the participants' perception of 'Green' (A) and 'Grey' (B) in three surveys (Cen reality, Cen online, Cen campaign) assessing the same locations; R – 'Cen reality', O – 'Cen online', C – 'Cen campaign' (the same pictures rated by different people online); each colour represents one of the three locations.

Tab. 4: Percentage agreements of the participants' perception of the variable 'Green' and 'Grey' and the algorithm's output based on the two following categories: vegetation, urban elements.

Variables	Cen reality	Cen online	Cen campaign	Ove campaign
Green	0.83	0.65	0.54	0.53
Gray	0.82	0.88	0.75	0.81

The participant's perception of an urban environment ('Grey') matched the algorithm's output better than their perception of a natural-like environment ('Green'). The percentage agreements were higher in the online surveys, with the highest score of 0.88 for *Cen online*. All results in Tab. 4 show a 50% agreement at the lowest. Additional results of the Kappa coefficient can be found in Tab. 10 in Appendix 5 – Image mining algorithm.

3.6 Reliability of the paid campaign

3.6.1 Cen online vs Cen campaign

This segment tries to answer if and how the results of two online surveys (*Cen online*, *Cen campaign*) varied because they comprised different people rating the same pictures. The following nine variables showed a significant difference between the two data sets ($p<0.05$): 'Maintenance', 'Sound', 'Irritated', 'Restricted', 'Threatened', 'Rejected', 'Anxious', 'Water', and 'Cared For' with a bordering p-value. These differences can be explained by the wider spread of the *Cen campaign* data points that may be contributed to the larger sample size or that those participants comprised a more heterogenous group varying in their answers. The IQR's of *Cen campaign* were generally high, mostly close to 2. The same observation of high the IQR was seen for the second data set from the paid campaign (*Ove campaign*).
Another point of view relies on plotting all data points of the perceptions-feelings relationships, i.e. the correlation coefficients in the real-life setting against the same correlations within *Cen online* and *Cen campaign*, resulting in similar distributions as seen in Fig. 9 below. Furthermore, additional text with two box plots in Appendix 4 speaks to the indecisive nature of the statistical differences.

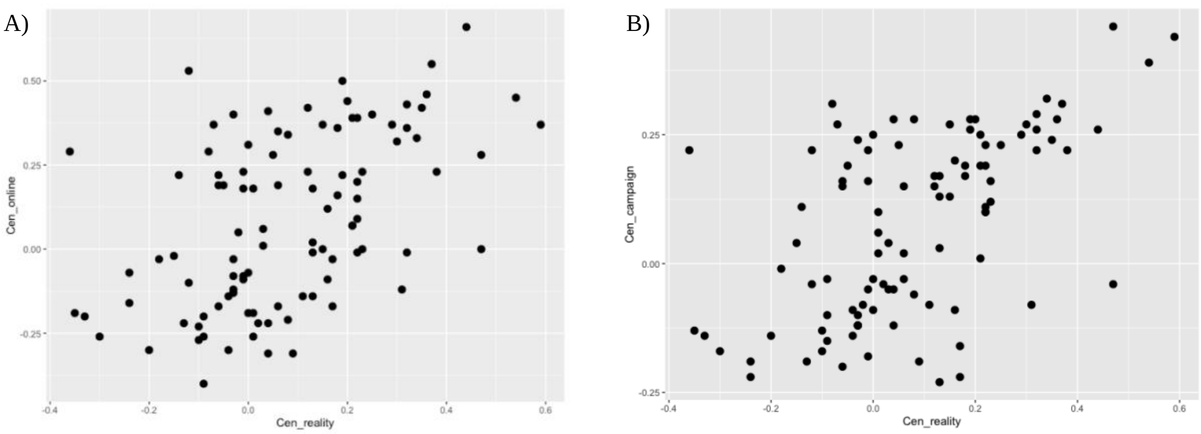

Fig. 9: The distribution of the feelings-perceptions relationships for the two online surveys assessing the same locations, but by different people, plotted against the same correlations from the real-life setting; A) Cen reality (x-axis) against Cen online, B) Cen reality (x-axis) against Cen campaign.

3.6.2 Cen campaign and Ove campaign

The perception of a water-dominated environment assessed by the participants from the paid campaign showed that these participants examined our questions because of an apparent difference in their ratings for the picture that captured a visible water body. However, some participants answered as if the other pictures would represent a water-dominated environment, though they could not observe it in them. To some extent, this phenomenon can be seen in *Cen online* as well, which leaves room for doubts about the reliability of these responses and, in general, suggests a careful examination of such surveys in future studies.

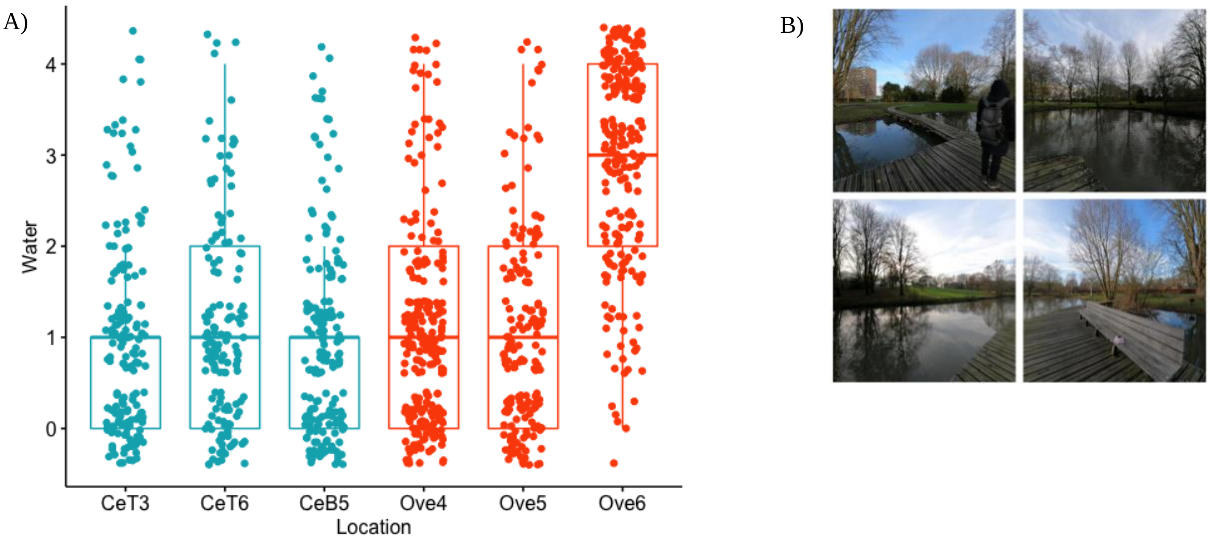

Fig. 10: Relevance of the paid campaign data – A) The box plot shows the participants' perception of a water-dominated environment in the two online surveys from the paid campaign assessing different areas of Utrecht, each including three locations; Cen campaign in blue, Ove campaign in red. B) The Ove6 location - the only location with a visible water body.

4 Discussion

Positive feelings had the highest ratings for all locations across four examined data sets, while the negative ones seemed to play a minor role in people's judgment. In terms of the perceptions, 'Social Connection', 'Uniqueness' and 'Entertainment' were rated higher in the real-life setting than online. Other perceptional variables worth mentioning were 'Green', 'Colours', and 'Sound' because of their clustering effect with the most pronounced feelings. In the data from the tapresearch campaign, these perception-feelings relationships had consistently lower correlation coefficients due to a larger IQR, which was the biggest observable difference across all four data sets. We were able to find significant correlations between the participants' perception of a natural-like and urban environment (the variables 'Green' and 'Grey'), and the effect of those properties on negative feelings as 'Green' showed negative correlations with the feelings 'Irritated' and 'Restricted' in contrary with 'Grey' that was positively correlated with the two feelings. Moreover, those perceptions showed high agreement with the objective measurements processed by the deep learning algorithm.

4.1 Strengths

Not many other studies have been done on this subject. Similar studies have been either conducted in laboratory settings or by researchers themselves (19,20) and not like in our case by sending out an open invitation to the online survey and recruiting participants by other means. Our strategy allows targeting people nationwide, possibly gathering data from different demographic groups, and makes it easy to do such surveys at scale. The survey was based on the previously developed paper version of a questionnaire created explicitly to assess perceived environments in the urban landscape. This questionnaire was adapted to the online environment and had undergone several testing rounds before publishing the online version. For example, measuring responders' duration resulted in our decision to include three locations in the survey which took about 10 minutes to complete. Answering the questions in the real-life setting was slightly quicker, but the total duration made up for 30 minutes because of the necessity of moving around to other locations. Several strategies were implemented during the participants recruitment, including the Tapresearch campaign, that let us test their efficiency and feasibility. This campaign allowed us to study variability and overarching patterns in the answers from different groups of people with a larger sample size. Even with the limited number of locations, we selected different environmental types, which reflected how the participants perceived a natural-like and urban environment (Fig. 8).

4.1.1 Algorithm's output

Besides the main result of high agreements between the participants' perception of a natural-like and urban environment and the algorithm's output, one striking finding was the algorithm's calculations of the amount of 'vegetation' for the same picture with- without leaves that resulted in almost exact percentages. In other words, the algorithm recognised trees as trees even without foliage, which is a desirable feature as people also perceive nature this way. However, that might also be the reason for the lower agreements of natural environments online (see Tab. 4 with the percentage agreements for 'Green'), where some pictures of vegetation appeared without leaves as they were taken in wintertime. The panels with picture overlays (Fig. 14) demonstrate this finding.

4.2 Limitations

As a pilot study, we encountered many obstacles during the study design and setting up the experiment, especially combining the assessment in the real-life setting and the online survey and or creating a single-entry point for all survey versions. The main limitation was the small number of participants (N=22) for the real-life setting. Ideally, the initial plan was to have, at least, a similar number of participants for the other two areas, which would make it possible to compare people's perceptions from the typologically same environments, characterised by the amount of 'vegetation' and 'urban elements', yet plausibly inducing different feelings in people. Finding enough participants willing to do both the assessment in the real-life setting and online survey was challenging. In general, choosing enough representative sites with specific characteristics is usually not performed due to feasibility (14, pp. 571). Another limitation in the design was the temporal discrepancy of the pictures used in the online surveys, as they were taken at the end of 2020, and the assessment in the real-life setting conducted in summer 2021. A change in the dynamics of the environment can happen suddenly, sometimes only within a couple of hours, and as a consequence, this type of temporal discrepancy is almost impossible to mitigate in practice. The second phase of the study may have been affected by some confounding factors that were not considered, e.g. weather conditions or the time of day. On a technical note, the layout in Qualtrics produced a small picture size with no option to enlarge the images, hence it might have been difficult for the participants to assess them as well as they could.

4.2.1 Result interpretation

Another drawback of the limited number of locations can be noticed through the variables that showed high positive correlations in the Reality vs Online comparison across all three locations, specifically the two negative feelings 'Threatened' and 'Rejected'. These feelings were neither induced at the locations nor online. It would be valuable to assess more locations to see why this occurred, especially outside a city environment or in neighbourhoods of low maintenance. The variable 'Water', representing a water-dominated environment, was the last of the three variables that showed those positive correlations in the same comparison, but the only one that also showed a significant difference between the two settings. Nevertheless, its

correlation for the CeB5 location had a significant high coefficient (0.63 ($p<0.002$)), despite the mean value for 'Water' was three times higher in the real-life setting than online. Therefore, some of these correlations may not provide an accurate picture of the actual conditions. However, this variable became necessary when we assessed the reliability of the paid campaign data by investigating its perception by the participants from *Cen campaign* and *Ove campaign* (Fig. 10). Lastly, the clustering effect of the feelings-perceptions relationships should be considered when designing future studies, as online surveys only based on pictures might result in homogeneous ratings that could uniform potential contrast across studied environments.
Somewhat misleading results related to this issue occurred for 'Cared For'. In two online surveys (*Cen online* and *Cen campaign*), its median was twice as large for most of the locations as opposed to the real-life setting. No sufficient explaining this result has been found.

4.3 Other studies

4.3.1 People's feelings towards greenery

Where the affinity of positive emotional states to being in a forest, park, or generally outside in natural environments began is unknown to us. Many attempts to explain this phenomenon through a lens of evolutionary psychology have failed to fully answer a simple question of why humans are drawn to nature (22–26). One thing is a theoretical framework, but denying actual results is difficult. For example, the feeling of being free had the highest score; its mean was 3.32, at the location with the most vegetation in the real-life setting. Only this location gained higher ratings of being free in the real-life setting as the other two had a lower rating of this feeling consistently in the real-life setting compared to the online. The other two feelings (Stimulated, Calm) with the highest ratings were again related to the greenest location in the real-life setting. That said, a study on the restorative effect of natural and non-natural environments has found that the restorative properties of being away only increased in the natural surroundings in the real-life setting but not in the built environment, which underpins our findings (20).
Similarly, data from a survey of about 20,000 participants on well-being has shown that the participants felt better in outdoor natural environments than urban environments (27). Results from another study of stress recovery in virtual environments have shown a positive relationship between the observation of green spaces and stress reduction. In contrast, urban environments did not produce the same results (15). Vegetation elements in our study may have played the same role by elevating positive feelings in people. However, not all green spaces constantly evoke positive feelings in people, demonstrated by our results from the CeT6 location. That is why a significant role in people's perception of greenary might play its quality or diversity that can contribute to the overall experience (28).
Our findings lead to a similar conclusion regarding safety as their interpretation is not entirely coherent with the corresponding literature. Our participants felt safer at the greenest location in the real-life setting than online, but the perception of safety was generally similar between the two settings and the locations with a different amount of vegetation. It has been shown that dense green spaces can negatively affect perceived safety in virtual reality (7). A common assumption that vegetation evokes lesser safe places in people can be neither proven nor discarded here, especially without a prior intention to study this effect in the study design. The literature rather points to other factors like the context, vegetation density and structure, and whether we talk about urban or rural areas as more comprehensive explanations (29,30).

4.3.2 Lower reliability for social aspects of the environment from pictures

Social Connection was the only variable that had no significant correlation in the Reality vs Online comparison among the three locations and showed a significant difference between the two settings at the same time. The participants consistently rated their surroundings as more socially engaging in the real-life setting than online. Other studies assessing the use of Google Street View for virtual auditing of urban environments support our findings. They generally point out that aesthetics or social characteristics had lower correlations coefficients, i.e. agreements, than other categories (19,31,32). In two of these studies, the authors elaborate on this issue by explaining possible causes with reasons such as a low social activity observed in certain pictures that might have been taken in the early morning or due to nonvisible parts of the scene covered by vehicles. However, Badland et al. contradict this view on aesthetics with their results of a perfect to an acceptable agreement, although they did not explain what the label 'Aesthetics' meant to be in their study (33). These auditing studies looked at more objective measures of the environment with no interest in subjective experience. In some cases, the aesthetics category was combined with social aspects of the environment and other qualities, which would be categorised as 'Maintenance' or 'Organisation' in our case.

This study imposed a different issue – the pictures were taken in the autumn and winter of 2020 during the Covid-19 pandemic, possibly leading to a changed perception of the social aspect of the environment by the participants, even unconsciously. Another contributing factor may have been the relatively small size of the pictures in the online survey. On the one hand, these drawbacks of the design might have resulted in lower ratings of 'Social Connection', 'Uniqueness', and 'Entertainment' online, but on the other hand, we did not find a big difference between the real-life setting and online for 'Maintainance' or 'Organisation'. Furthermore, the CeT3 location, a narrow street in the city centre, has been redesigned with wider sidewalks and limited traffic due to the new design being finished at the end of 2020 (34). When we consider this circumstance, it seems logical that the participants perceived this location as more engaging, especially in terms of social connection and entertainment. This observation corresponds with the participants feeling less calm at the location than online. The assessment in the real-life setting was conducted in the summertime, with many bars open and more pedestrians and cyclists passing by.

4.3.3 Auditory sense

The online survey was purposely created using panoramic pictures but lacked other stimuli. The strongest relationship between the participants' feelings and their perceptions in the real-life setting was 'Sound-Calm'. Further, the two surveys assessing the same pictures showed inconsistency in the participants' perception of sound in the online environment (Fig. 13). Nonetheless, the results demostrate that people can broadly sense the soundscape evoked by pictures without the actual auditory experience. Other studies that looked into stress reduction and mood changes in a virtual nature-evoking environment have found a positive influence of sound compared to only using a visual stimulus (35,36). Our finding supports the importance of the auditory sense in studies on perceived environments.

4.4 The big picture

4.4.1 The case against the real-life assessment

This project aimed to answer whether online surveys are reliable tools in researching perceived environments. Instead of only looking at the differences and similarities between the real-life setting and online, it may be helpful to view this issue through possible confounding factors that can occur during on-site assessments and might be mitigated by the online alternative. Feasibility plays a crucial role in addressing this issue. It can be argued that as one judges a scene by looking at a picture and then doing the on-site assessment, such a person will be affected differently due to multiple ever-changing factors. One can ask what would happen if this person went and assessed the place another time, meaning in a different season or weather conditions, time of a day, or being tired instead of excited. Such assessments contain several confounding factors that are difficult to be moderated. Thus, future studies should disentangle to which degree those factors impact the results.

Nevertheless, some of these conditions may be possible to control in the online environment. A study, which looked at the perception of street vendors, has found significantly lower levels of annoyance, in other words, better acceptance of those vendors at night than in daytime (37). Preferences expressed towards the same scene can be further altered by lighting conditions and an added or erased sound, as the results of a study using a computer simulation have suggested (38). These examples show that the same scene can evoke a change in perception under different circumstances. When conducted correctly, online surveys can play an important role to tackle these challenges.

4.4.2 Subjective experience vs objective properties of a place

Our perception of the environment is undoubtedly tight to its physical properties. However, we combine these stimuli with cognitive processes that are, in part, based on a person's attitude toward a place, biases linked to other experiences from similar places, possibly place attachments (39,40). To distinguish between the two approaches and address the causal attributes of the environment by which someone is affected can be a challenging task. Then we can ask to what extent is using quantifiable and objective methods of assessing people's perceptions of their surroundings realistic. Still, we have presented some positive results regarding the notion of people's perceptions corresponding with the object recognition by the algorithm. Cognitive neuroscience research has pointed out that our perception depends on one's goals, integrates initial assumptions with the stimuli received by our senses, resulting in a probabilistic outcome, and stretches sometimes forgotten consequences of the difference between our judgement and perceptions (41,42). However, this theoretical background describing the ambiguity of our everyday experiences seems impossible to implement into the current design of other types of studies. Thus, quantifying people's perceptions based on objective attributes of the environment in the real-life setting may take a long time, but this project sheds light on some possible strategies.

4.5 Opportunities

The topic of perceived environments can be studied by multiple techniques and approaches. In the scope of online surveys and based on our findings and other studies, we suggest implementing auditory experience into the online surveys so that people would have a more realistic understanding of the scene they are assessing. The question is how to accurately record the soundscape of a place that will be representative and not biased by the researcher's assumptions about that place. A larger sample size of sounds always taken at the same time of day with randomised selection may be helpful here or preset participants with multiple pictures and related sounds of the same place on different occasions, e.g. rush hours, on the weekend night time. The randomised selection strategy can also be used for pictures depending on the amount of vegetation and urban elements predefined by the algorithm and other criteria. An even more appropriate solution may be to take a short clip of a scene instead of using still pictures. These suggestions come with two obstacles: 1) ethical concerns of not bothering people with the camera, 2) using the algorithm for analysing the footage may be technically challenging. A new procedure would have to be developed, making the process more complex. Still, utilising the algorithm further is, without a doubt, another opportunity for the future.

With a larger sample size and narrower scientific interests, the algorithm's output could contribute to a better understanding of whether objective attributes of the environment are mirrored in people's perceptions of a place. Lastly, we argue that using online surveys for studying perceived environments can, regarding the primary research question of this study, paradoxically reduce some confounding factors by providing a controlled environment without ever-changing weather conditions and other external factors of the real environment. Above all, this approach can reduce time and cost for both the researcher and the participants. Different technical solutions of how to incorporate objectively measured conditions of the environment can be implemented in the future, for example, some other studies with the interest of assessing urban spaces and natural environments have used methods like Isovist properties, spatial frequencies, or suggesting to investigate fractal structures (43–46). For this type of assessment, another opportunity is implementing virtual reality that could simulate a city tour and possibly allow to study people's perceptions on both the subjective and objective level, e.g. analysing answers from a questionnaire and measuring physiological signals in people (10,11).

5 Conclusion

This project questioned a widely accepted assumption that online surveys and mainly the pictures used in them provide an accurate representation of the actual environment to study people's perceptions of it. The online survey was created of three locations in the city of Utrecht, and its results were compared to the assessment in the real-life setting of those places by the same participants. The survey was focused on assessing people's feelings and subjective perceptions, comparing 20 numeric variables in total. The deep learning algorithm then analysed photographs of the locations for object recognition. In that way, it was possible to compare objective properties of the environment, calculated by the algorithm's output with perceptual variables from the participants. Additional online responses were gathered through a paid campaign using novel recruiting strategies. Those data helped with the result interpretation and answered the secondary aim of this study.

The main findings have shown that most variables did not differ between the two settings. Those were primarily positive and negative feelings with high and low ratings, respectively. Some variables, namely 'Entertainment', 'Uniqueness', 'Social Connection', were appreciated by the participants consistently less in the online surveys showing that the pictures lacked certain attributes of the environment, unable to evoke a realistic image of our locations. On the other hand, the perception of 'vegetation' and 'urban elements' (the variables 'Green' and 'Grey') did not change between the real-life setting and online and even matched with the algorithm's output. All agreements, calculated for the two variables from four surveys, were consistently above 50%, with the highest agreements for 'urban elements' online. These findings demonstrate a promising step to link people's subjective perception of various environmental attributes with objective measurements. The data provided by the paid campaign combined with the answers from our main participants point to a clustering effect of certain perceptions and feelings, although some of them conflict with the real-life setting. Of these relationships, sound perception appeared to be significant for the participants in the real-life setting but was inconsistent in their perception online at the same time. Thus, we propose implementing the auditory sense into the online environment, ideally by including a video of a scene.

To our knowledge, this was the first study approaching the topic of perceived environments by using an online survey and comparing that to the real-life setting and further integrating the object recognition algorithm. The project would benefit from a larger sample size of real-life participants and more assessed sites. Another challenge was interpreting the results, as little has been done in this research area. Despite that, the project has shown the potential of using online surveys to study perceived environments and paved the way for future attempts to tackle this issue. Future studies may focus on addressing potential group differences in people's perceptions, making a specific use case for utilising the algorithm further or looking into the added value of soundscape in the online environment.

6 Appendices

6.1 Appendix 1 – Demographics and descriptives

Tab. 5: Demographic information that includes two the online campaigns, i.e. the table represents three different groups of participants.

		Cen reality / Cen online		Cen campaign		Ove campaign	
Groups	**Categories**	**count**	**prop [%]**	**count**	**prop [%]**	**count**	**prop [%]**
Gender	**Male**	17	77.27	58	34.12	156	68.12
	Female	5	22.73	111	65.29	72	31.44
	NA	0	0	1	0.59	1	0.44
Age	**18 - 30**	16	72.73	110	64.71	45	19.65
	31 - 50	4	18.18	40	23.53	163	71.18
	51 - 65	2	9.09	15	8.82	17	7.42
	66 - above	0	0	5	2.94	4	1.75
Education	**High School**	1	4.55	52	30.59	72	31.44
	MBO	2	9.09	51	30	80	34.93
	HBO	0	0	43	25.29	51	22.27
	University	19	86.36	24	14.12	26	11.35
Childhood in NL	**Yes**	8	36.36	159	93.53	213	93.01
	No	14	63.64	6	3.53	12	5.24
	Partly	0	0	5	2.94	4	1.75
Living in Utrecht	**Currently living**	17	77.27	10	5.88	19	8.3
	Yes, in the past	3	13.64	27	15.88	30	13.1
	No, never	2	9.09	129	75.88	166	72.49
	Born and raised	0	0	4	2.35	14	6.11

When the data on the real-life participants were compared to the two data sets of the paid campaign, some differences can be found in every category. *Cen campaign* has the opposite ratio of sex (65.29% for females), *Ove campaign* is the only data set with most participants in the age group of 31-50. Both *Cen campaign* and *Ove campaign* have a percentage of university graduates below 15 percent compared to the real-life participants, exceeding 85%. Another noticeable difference is the participants' residency outside Utrecht in most cases for the online campaigns. On the other hand, over 90% of those participants were born or raised in the Netherlands.

Tab. 6: A table showing descriptive statistics on perceptual variables and feelings for four surveys; Cen reality, Cen online, Cen campaign assessing the same locations (the Cen area) and Ove campaign assessing the Ove area.

		Cen reality							Cen online							Cen campaign							Ove campaign						
Variables		count	mean	sd	p0.25	median	p0.75	IQR	count	mean	sd	p0.25	median	p0.75	IQR	count	mean	sd	p0.25	median	p0.75	IQR	count	mean	sd	p0.25	median	p0.75	IQR
Peceptions	Maintenance	66	0.95	1.00	1	1	2	1	66	1.17	0.78	1	1	2	1	510	0.81	0.98	1	1	1	0	684	0.70	1.02	0	1	1	1
	Organization	66	0.94	0.97	1	1	2	1	66	1.05	0.88	1	1	2	1	510	0.82	0.98	1	1	1	0	684	0.76	0.99	0	1	1	1
	Social Connection	66	0.97	0.86	1	1	1	0	66	0.58	0.98	0	1	1	1	510	0.40	1.12	-1	1	1	2	684	0.32	1.16	-1	1	1	2
	Entertainment	66	0.82	0.98	0	1	1	1	66	0.34	1.04	0	1	1	1	510	0.11	1.21	-1	0	1	2	684	-0.02	1.23	-1	0	1	2
	Uniqueness	66	0.65	0.95	0	1	1	1	66	0.02	0.99	-1	0	1	2	510	0.01	1.17	-1	0	1	2	684	0.04	1.26	-1	0	1	2
	Colours	66	0.91	0.89	1	1	1	0	66	0.49	1.02	0	1	1	1	510	0.48	1.11	0	1	1	1	684	0.63	1.12	0	1	1	1
	Sound	66	0.71	1.13	0	1	1.75	1.75	66	0.86	0.92	0	1	1	1	510	0.53	1.12	0	1	1	1	684	0.71	1.06	0	1	1	1
	Green	66	2.18	1.41	1	2	3	2	66	1.98	1.40	1	2	3	2	510	2.19	1.37	1	2	3	2	684	2.48	1.12	2	3	3	1
	Gray	66	1.95	1.28	1	2	3	2	66	1.86	1.28	1	2	3	2	510	1.99	1.24	1	2	3	2	684	1.55	1.04	1	1	2	1
	Water	66	0.73	1.05	0	0	1	1	66	0.43	0.83	0	0	1	1	510	1.00	1.08	0	1	2	2	684	1.77	1.44	1	1	3	2
Feelings	Safe	66	2.58	1.08	2	3	3	1	66	2.26	1.15	2	2	3	1	510	2.26	1.08	2	2	3	1	684	2.13	1.25	1	2	3	2
	Stimulated	66	1.83	1.13	1	2	2	1	66	1.62	1.08	1	2	3	2	510	1.68	1.20	1	2	3	2	684	1.58	1.29	0	1	3	3
	Cared For	66	0.88	1.05	0	1	1.75	1.75	66	1.44	1.15	0	2	2	2	510	1.78	1.18	1	2	3	2	684	1.80	1.33	1	2	3	2
	Calm	66	2.23	1.17	1	2	3	2	66	2.42	1.07	2	2.5	3	1	510	2.45	1.18	2	3	3	1	684	2.58	1.27	2	3	4	2
	Free	66	2.48	1.10	2	3	3	1	66	2.50	1.06	2	3	3	1	510	2.52	1.19	2	3	4	2	684	2.47	1.31	2	3	4	2
	Irritated	66	0.70	1.01	0	0	1	1	66	0.55	0.77	0	0	1	1	510	0.99	1.05	0	1	2	2	684	0.85	1.12	0	0	1.25	1.25
	Restricted	66	0.71	0.99	0	0	1	1	66	0.80	1.06	0	0	1.75	1.75	510	1.25	1.10	0	1	2	2	684	1.07	1.15	0	1	2	2
	Threatened	66	0.48	0.81	0	0	1	1	66	0.52	0.83	0	0	1	1	510	0.99	1.08	0	1	2	2	684	0.91	1.16	0	0	2	2
	Rejected	66	0.35	0.71	0	0	0	0	66	0.30	0.61	0	0	0	0	510	0.93	1.09	0	1	2	2	684	0.80	1.11	0	0	1	1
	Anxious	66	0.45	0.75	0	0	1	1	66	0.53	0.88	0	0	1	1	510	1.03	1.10	0	1	2	2	684	0.91	1.18	0	0	2	2

Notably, the means of the three highest-rated feelings belonged to the real-life setting at the same location – the CeB5 location (Free=3.32, Calm=3.04, Safe=2.86). Furthermore, only this location was characterised by having the highest and lowest mean values in the real-life setting. The CeT6 location, on the other hand, showed the opposite trend, where the three lowest mean values came from the pictures, and so did the two most pronounced feelings. Negative feelings gained low ratings across all data sets. However, especially those in *Cen campaign* had, on average, higher values than in *Cen reality* and *Cen online*, for example, 'Rejected' as the least manifested with the mean of 0.93.

As for the variables representing perceptions, 'Uniqueness' and 'Entertainment' were the only ones that showed negative mean values, speaking to the ordinariness and boringness of the scenes at CeT3 and CeT6 in both online surveys (*Cen online* and *Cen campaign*), unshown data. The same can be seen in *Ove campaign* (the data set with different pictures), unshown data. Interestingly, the same variables in the real-life setting showed, on average, higher scores, especially 'Uniqueness' with the mean of 0.65, whereas in the remaining online data sets, it averaged just above zero.

The last difference between the results of the real-life participants and those from the online campaign worth pointing out lies in a higher IQRs of *Cen campaign* (N=510), supported by the same observation in *Ove campaign* (N=684), compared to the other data sets (*Cen online* and *Cen reality*).

6.2 Appendix 2 – Reality vs Online

Tab. 7: Net values represent the absolute differences between the participants' ratings in the real-life setting and online, i.e. a net positive means higher ratings in the real-life setting.

		All locations			CeT3 location			CeT6 location			CeB5 location		
Variables		**mean**	**sd.**	**count**	**mean**	**sd.**	**count**	**mean**	**sd.**	**count**	**mean**	**sd.**	**count**
Perceptions	**Maintenance**	-0.2	1.2	66	-0.3	1.1	22	-0.7	1.3	22	0.4	0.8	22
	Organization	-0.1	1.2	66	-0.3	1.0	22	-0.4	1.5	22	0.4	1.0	22
	Social Connection	0.4	1.3	66	0.8	1.6	22	0.1	1.2	22	0.3	1.0	22
	Entertainment	0.5	1.3	64	1.0	1.3	22	0.2	1.5	20	0.2	1.1	22
	Uniqueness	0.6	1.0	57	0.6	1.3	19	0.4	0.9	17	0.8	0.9	21
	Colours	0.4	1.2	65	0.0	0.8	22	0.3	1.3	21	0.9	1.2	22
	Sound	-0.1	1.1	65	0.0	1.0	22	-0.4	1.5	21	-0.0	0.9	22
	Green	0.2	1.1	66	0.2	0.9	22	-0.3	0.8	22	0.7	1.2	22
	Gray	0.1	1.1	64	-0.3	1.4	22	0.6	0.9	20	0.2	0.9	22
	Water	0.3	0.9	61	-0.2	0.7	22	0.2	0.6	17	0.9	0.9	22
Feelings	**Safe**	0.3	1.3	66	-0.2	1.1	22	0.2	1.1	22	1.0	1.3	22
	Stimulated	0.2	1.3	66	0.3	1.2	22	0.0	1.3	22	0.3	1.4	22
	Cared For	-0.6	1.4	66	0.0	1.0	22	-0.9	1.5	22	-0.9	1.4	22
	Calm	-0.2	1.6	66	-0.5	1.1	22	-0.9	1.4	22	0.9	1.8	22
	Free	-0.0	1.4	66	-0.2	1.1	22	-0.5	1.5	22	0.7	1.3	22
	Irritated	0.2	1.1	66	0.0	1.0	22	0.7	1.3	22	-0.3	0.6	22
	Restricted	-0.1	1.2	66	-0.3	1.4	22	0.1	1.2	22	-0.1	0.9	22
	Threatened	-0.0	0.9	66	-0.2	0.7	22	0.4	1.0	22	-0.3	0.8	22
	Rejected	0.0	0.5	66	-0.1	0.4	22	0.3	0.6	22	-0.0	0.5	22
	Anxious	-0.1	1.1	66	-0.2	0.7	22	0.1	0.8	22	-0.1	1.5	22

6.2.1 Results of the Wilcoxon test

The sample of 22 participants who did both assessments showed significant differences ($p < 0.05$) for six (out of 20) following variables: 'Social Connection', 'Entertainment', 'Uniqueness', 'Colours', 'Water', and 'Cared For' showing the highest $W = 2780$ ($p=0.004$).

6.2.2 Differences between the Cen locations within three data sets

In terms of the differences between the locations, two variables stood out, namely 'Green' and 'Grey' in all three data sets (*Cen reality*, *Cen online*, *Cen campaign*), and a few more variables with high chi-squared values appeared in *Cen campaign*: 'Social Connection', 'Sound', 'Free', 'Calm'. The post-hoc analysis done for 'Green' and 'Grey' showed a difference across all three locations in each data set despite one case in *Cen online*, the combination of the CeT6-CeB5 locations.

Tab. 8: The Spearman correlations of the ratings in the real-life setting and online by the same participants (N=22); numbers in bold indicate high correlations across all three locations; p-values with '' mean p < 0.000.*

		CeT3		CeT6		CeB5	
Variables		**estimate**	p.value	**estimate**	p.value	**estimate**	p.value
Perceptions	**Maintenance**	0.18	0.415	-0.14	0.544	0.51	0.015
	Organization	0.30	0.172	0.07	0.767	0.45	0.034
	SocialConnection	-0.30	0.168	-0.12	0.597	-0.06	0.789
	Entertainment	-0.31	0.159	0.24	0.312	0.46	0.031
	Uniqueness	-0.03	0.905	0.65	0.005	0.53	0.014
	Colours	0.32	0.144	0.10	0.656	0.33	0.131
	Sound	0.50	0.018	-0.09	0.712	0.26	0.240
	Green	0.25	0.266	0.55	0.009	0.10	0.643
	Gray	0.14	0.542	0.43	0.060	0.58	0.004
	Water	**0.76**	0.000*	**0.52**	0.033	**0.63**	0.002
Feelings	**Safe**	0.22	0.319	0.48	0.024	0.44	0.039
	Stimulated	0.24	0.287	0.25	0.260	0.36	0.099
	CaredFor	0.49	0.020	0.11	0.626	0.17	0.461
	Calm	0.43	0.046	-0.16	0.468	-0.29	0.193
	Free	0.35	0.107	-0.05	0.814	0.00	0.990
	Irritated	0.24	0.285	0.24	0.274	0.29	0.190
	Restricted	0.06	0.807	0.41	0.061	0.59	0.004
	Threatened	**0.80**	0.000*	**0.48**	0.024	**0.62**	0.002
	Rejected	**0.63**	0.002	**0.63**	0.002	**0.78**	0.000*
	Anxious	0.59	0.004	0.32	0.153	0.14	0.545

The variables 'Threatened', 'Rejected', and 'Water' showed significant positive correlations between the online and the real-life setting for each location, while there was no significant correlation for the variables 'Social Connection', 'Colours', 'Irritated', 'Stimulated', and 'Free'. The highest estimate was for 'Threatened' at CeT3 was 0.80 (p=0.0001). The other three variables, namely 'Uniqueness', 'Safe', 'Grey', showed significant positive correlations for two out of the three locations. Overall, no negative correlation has been found with a significant p-value.

6.3 Appendix 3 – Correlation structures

Tab. 9: The Spearman correlations showing the relationships of the perceptual variables with feelings in four surveys; p-values are not shown, but as for Cen reality and Cen online the correlation coefficients of about 0.25 and higher were statistically significant (p<0.05) and in the two online campaings only coefficients lower than 0.10 were not signigicant.

Variables		Cen reality	Cen online	Cen campaign	Ove campaign
Maintenance	Irritated	-0.10	-0.23	-0.17	-0.13
	Safe	0.38	0.23	0.22	0.24
	Restricted	0.11	-0.14	-0.08	-0.10
	Threatened	-0.10	-0.27	-0.13	-0.11
	Stimulated	0.21	0.07	0.01	0.19
	Cared For	0.23	0.23	0.12	0.22
	Rejected	-0.13	-0.22	-0.19	-0.13
	Calm	0.18	0.16	0.17	0.23
	Free	0.16	0.12	0.20	0.20
	Anxious	0.13	-0.14	-0.23	-0.12
Organisation	Irritated	-0.00	-0.19	-0.09	-0.16
	Safe	0.12	0.23	0.15	0.28
	Restricted	0.06	-0.17	-0.03	-0.16
	Threatened	0.08	-0.21	-0.06	-0.16
	Stimulated	0.03	0.06	-0.05	0.14
	Cared For	0.15	0.37	0.13	0.23
	Rejected	0.16	-0.09	-0.09	-0.12
	Calm	0.22	0.09	0.10	0.26
	Free	-0.14	0.22	0.11	0.22
	Anxious	0.31	-0.12	-0.08	-0.14
Social Connection	Irritated	-0.06	-0.17	-0.20	-0.13
	Safe	0.13	-0.01	0.17	0.28
	Restricted	-0.04	-0.14	-0.14	-0.12
	Threatened	-0.03	-0.13	-0.12	-0.12
	Stimulated	0.18	0.36	0.19	0.23
	Cared For	0.06	0.35	0.15	0.27
	Rejected	-0.01	-0.09	-0.18	-0.07
	Calm	-0.07	0.37	0.27	0.28
	Free	0.08	0.34	0.28	0.27
	Anxious	0.09	-0.31	-0.19	-0.15
Entertainment	Irritated	-0.09	-0.20	-0.03	-0.10
	Safe	0.23	0.00	0.16	0.31
	Restricted	0.01	-0.26	0.02	-0.10
	Threatened	-0.15	-0.02	0.04	-0.06
	Stimulated	0.37	0.55	0.31	0.32

	Cared For	0.04	0.41	0.28	0.35
	Rejected	-0.18	-0.03	-0.01	-0.05
	Calm	-0.05	0.19	0.19	0.33
	Free	0.21	0.39	0.25	0.28
	Anxious	0.04	-0.31	-0.05	-0.15
Uniqueness	Irritated	0.00	-0.07	-0.03	-0.10
	Safe	0.13	0.18	0.13	0.28
	Restricted	0.01	-0.19	0.06	-0.05
	Threatened	0.03	0.01	0.04	-0.04
	Stimulated	0.19	0.50	0.28	0.31
	Cared For	-0.03	0.40	0.24	0.31
	Rejected	0.06	0.19	0.02	-0.01
	Calm	0.22	0.20	0.23	0.31
	Free	0.22	0.39	0.19	0.25
	Anxious	0.02	-0.22	-0.04	-0.07
Colours	Irritated	-0.35	-0.19	-0.13	-0.17
	Safe	0.21	0.07	0.19	0.39
	Restricted	-0.03	-0.08	-0.12	-0.15
	Threatened	-0.01	-0.08	-0.05	-0.12
	Stimulated	0.44	0.66	0.26	0.27
	Cared For	-0.12	0.53	0.22	0.34
	Rejected	-0.03	-0.03	-0.12	-0.10
	Calm	0.30	0.32	0.27	0.36
	Free	0.20	0.44	0.28	0.37
	Anxious	0.04	-0.22	-0.12	-0.16
Sound	Irritated	-0.24	-0.16	-0.22	-0.22
	Safe	0.32	-0.01	0.22	0.24
	Restricted	-0.09	-0.26	-0.15	-0.19
	Threatened	-0.03	-0.12	-0.10	-0.18
	Stimulated	0.29	0.37	0.25	0.19
	Cared For	-0.08	0.29	0.31	0.22
	Rejected	0.17	-0.03	-0.16	-0.18
	Calm	0.59	0.37	0.44	0.35
	Free	0.36	0.46	0.28	0.30
	Anxious	0.17	-0.17	-0.22	-0.16
Green	Irritated	-0.33	-0.20	-0.14	-0.10
	Safe	0.22	-0.01	0.11	0.27
	Restricted	-0.20	-0.30	-0.14	-0.03
	Threatened	-0.12	-0.10	-0.04	-0.04
	Stimulated	0.32	0.43	0.26	0.21
	Cared For	-0.36	0.29	0.22	0.26

	Rejected	-0.02	0.05	-0.08	-0.00
	Calm	0.47	0.28	0.46	0.32
	Free	0.54	0.45	0.39	0.35
	Anxious	-0.09	-0.40	-0.10	-0.03
Grey	Irritated	0.25	0.40	0.23	0.19
	Safe	-0.06	0.22	0.15	-0.07
	Restricted	0.35	0.42	0.24	0.17
	Threatened	-0.06	0.19	0.16	0.14
	Stimulated	-0.04	-0.30	-0.09	-0.01
	Cared For	0.47	-0.00	-0.04	-0.15
	Rejected	-0.01	0.18	0.16	0.19
	Calm	-0.24	-0.07	-0.19	-0.12
	Free	-0.30	-0.26	-0.17	-0.11
	Anxious	0.12	0.42	0.17	0.24
Water	Irritated	0.00	0.31	0.25	0.03
	Safe	0.01	0.18	0.10	0.19
	Restricted	0.05	0.28	0.23	0.04
	Threatened	0.32	0.36	0.29	0.06
	Stimulated	0.15	0.00	0.27	0.22
	Cared For	-0.01	0.23	0.22	0.23
	Rejected	0.34	0.33	0.32	0.05
	Calm	0.13	0.02	0.03	0.24
	Free	0.22	0.15	0.10	0.20
	Anxious	0.19	0.22	0.26	0.06

We can see the same clustering effect for the feeling ‘Cared For’ online, with high positive correlations, but that never held in the real-life setting. The mismatch was the most pronounced for the perception of greenness (Green), with a negative correlation of -0.36 in the reality. In contrast, the perception of urban environments (Grey) implies the complete opposite, with a positive correlation of 0.47 in the reality, and then online, the rest is close to 0.

Similarly, this pattern of opposing trends in the reality can be seen for the ‘Entertainment-Calm’ relationship, and even more clearly for the feeling ‘Anxious’, which always had positive or no correlation with the perceptions, but negative ones online. The highest negative correlation was imposed by ‘Green-Anxious’ with the coefficient of -0.40 in the *Cen online* data set.

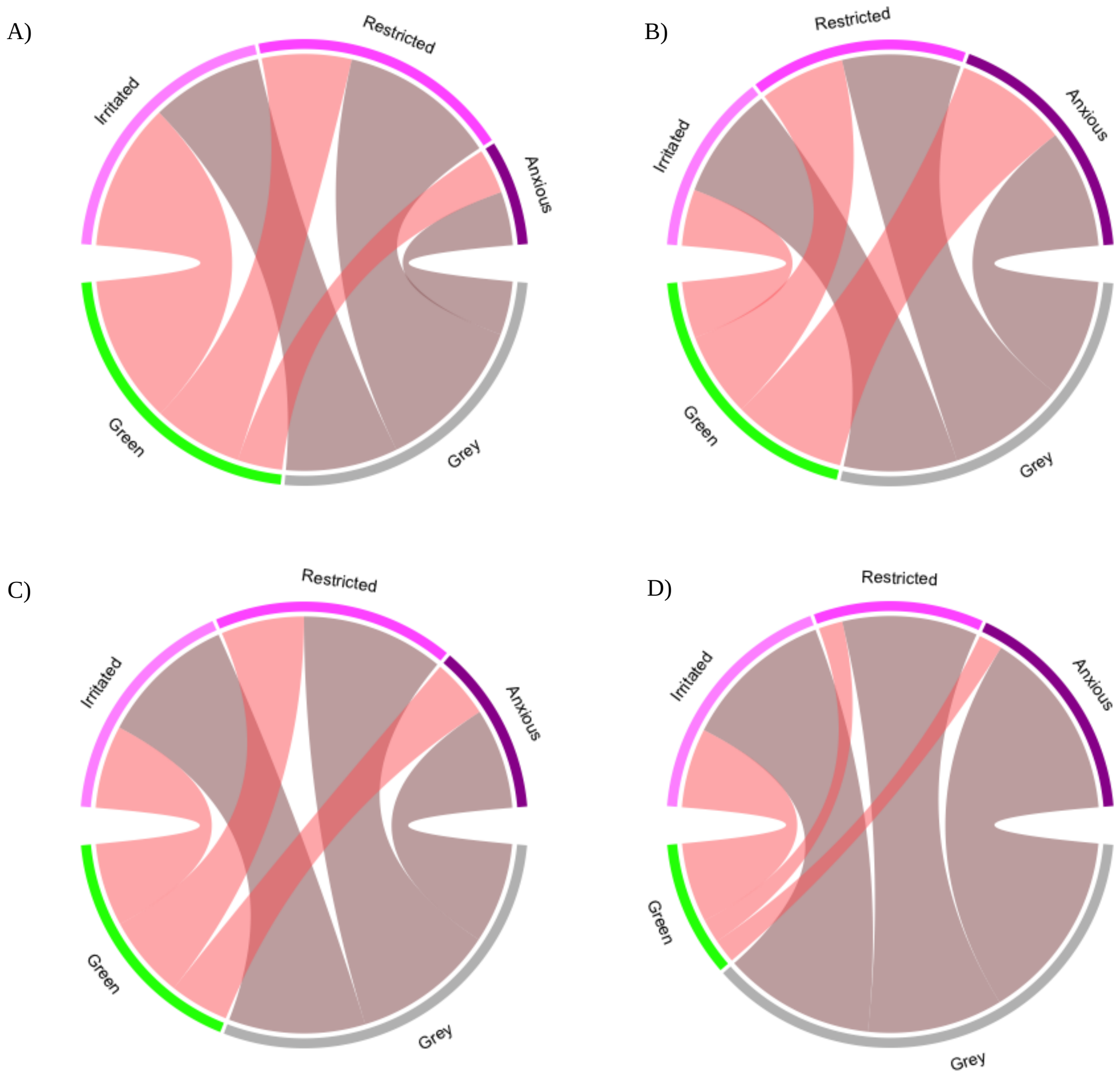

Fig. 11: Environmental types and negative feelings – the participants' perception of an urban ('Grey') and natural-like environment linked with negative feelings in the real-life setting and online; A) Cen reality, B) Cen online (the same participants online), C) Cen campaign (the same pictures rated by different people online), D) Ove campaign (different pictures rated by another group of people online); the width of the links indicates its strength, i.e. the correlation coefficient, and negative correlations are in red.

The patterns in Fig. 11 need to interpret with caution. Although the correlations of these relationships were significant and relatively high (Tab. 9), all negative feelings gained low ratings across all data sets. Therefore, the potential effect and relevance of the participants' perception of different environmental types on them might have been negligible.

6.4 Appendix 4 – Reliability of the paid campaign

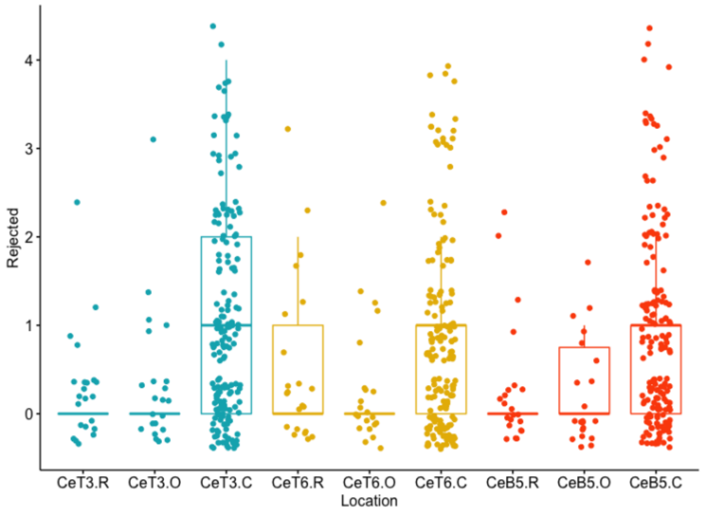

Fig. 12: 'Restricted' shows the same pattern in the ratings for three surveys (Cen reality, Cen online, Cen campaign), although this feeling was significantly different when Cen campaign was compared to Cen online. The same uplifted values can be seen for the CeT3 location across all the three data sets; R – 'reality', O – 'Cen online', C – 'Cen campaign' (the same pictures rated by different people); each colour represents one of the three locations.

'Restricted' showed a significant difference between *Cen online* and *Cen campaign*, but Fig. 12 illustrates that the ratings were similar across the three data sets, following the same pattern for all locations. This observation mostly applies to the other variables that were significantly different between *Cen online* and *Cen campaign* as well. Hence they do not disprove the relevance of *Cen campaign*, especially as this data set included a more significant number of

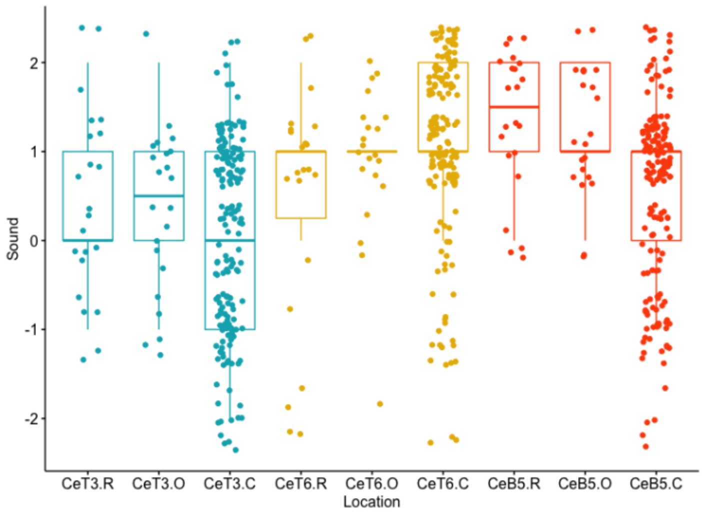

Fig. 13: The perception of 'Sound' varying in ratings for the CeT6 and CeB5 locations in Cen campaign compared to the perception by the main participants; R – 'reality', O – 'Cen online' (the same participants online), C – 'Cen campaign' (the same pictures rated by different people online); each colour represents one of the three locations..

participants (N=510, all three locations together), pointing to the inconclusiveness of these findings. For instance, 'Threatened' and 'Anxious' can be explained by their higher median (=1) in *Cen campaign*, whose IQR (=2) was twice as large as in *Cen online* or *Cen reality*.
In fact, 'Sound' (Fig. 13) is one of the few variables where *Cen campaign* shows a distinctly different perception with higher values for another location than the participants' ratings from *Cen reality* and *Cen online*.

6.5 Appendix 5 – Image mining algorithm

Tab. 10: The Kappa coefficient as an additional result to the percentage agreements of the participants' perceptions of 'Green' and 'Grey' and the algorithm's output for three data sets assessing the same locations in the real-life setting and online; p-values marked with '' showed p<0.000.*

Estimate	Cen reality	p-value	Cen online	p-value	Cen campaign	p-value
Green	0.66	0.000*	0.29	0.015	0.05	0.280
Gray	0.59	0.000*	0.72	0.000*	0.44	0.000*

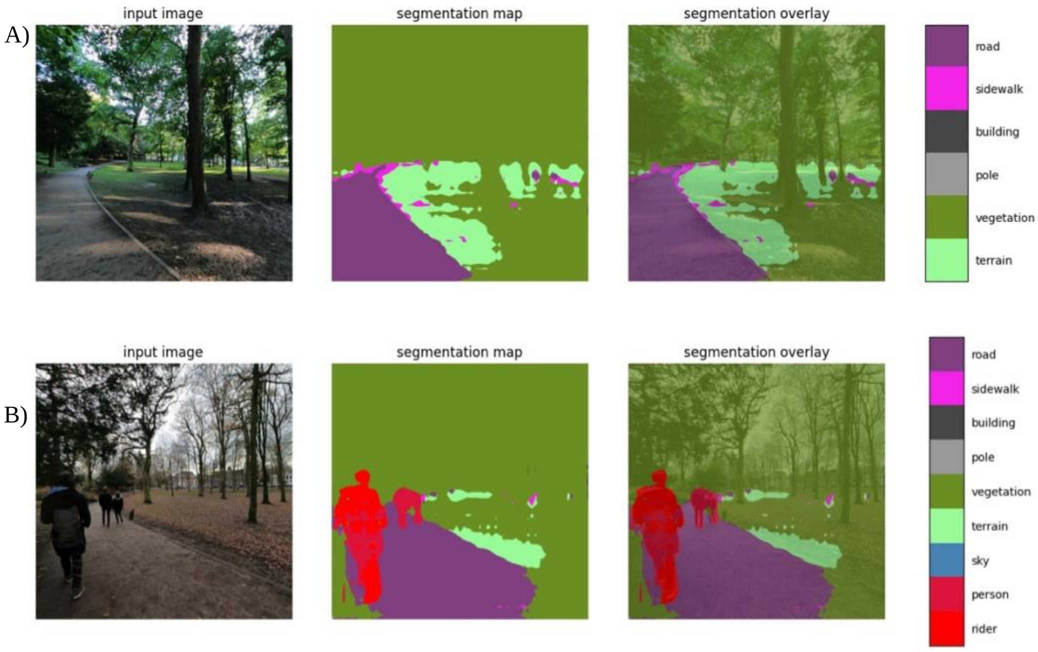

Fig. 14: An illustration of the same amount of vegetation calculated by the deep learning algorithm for the CeB5 location; A) A picture with a vegetation cover (summer 2021), B) A picture without a vegetation cover (wintertime).

6.6 Appendix 6 – Search for new locations

Part of the project was to expand and choose new locations representing missing environments. New locations would bring more diversity by including areas with a low population density, cutting-edge architecture, industrial zones, and agricultural fields. Further, a map with the current and new possible locations was created in Google Earth, and the new suggested locations can be seen listed below. They were chosen based on the condition of being close to residential areas.

1) Modern buildings – high-rise buildings, cutting-edge architecture
   a) Central station – the top square at the entrance, a modern public space
   b) Utrecht Science Park – modern and an easy target for the survey, very specific environment itself with context for the students who live there
2) Grey areas – warehouses around Ikea
3) Wild nature – Panbos (national forest) and Amelisweerd park near the Science Park Campus
4) Agricultural fields – near Fort de Gagel (castle)

References

1. Ulrich RS. View through a window may influence recovery from surgery. Science (80-). 1984;224(4647):420–1.
2. Twohig-Bennett C, Jones A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. Environ Res [Internet]. 2018;166(June):628–37. Available from: https://doi.org/10.1016/j.envres.2018.06.030
3. Villeneuve PJ, Jerrett M, G. Su J, Burnett RT, Chen H, Wheeler AJ, et al. A cohort study relating urban green space with mortality in Ontario, Canada. Environ Res [Internet]. 2012;115(2012):51–8. Available from: http://dx.doi.org/10.1016/j.envres.2012.03.003
4. De Vries S, van Dillen SME, Groenewegen PP, Spreeuwenberg P. Streetscape greenery and health: Stress, social cohesion and physical activity as mediators. Soc Sci Med [Internet]. 2013;94:26–33. Available from: http://dx.doi.org/10.1016/j.socscimed.2013.06.030
5. Kim M, Rupprecht CDD, Furuya K. Residents' perception of informal green space-a case study of Ichikawa City, Japan. Land. 2018;7(3):1–20.
6. Matsuoka RH, Kaplan R. People needs in the urban landscape: Analysis of Landscape And Urban Planning contributions. Landsc Urban Plan. 2008;84(1):7–19.
7. Tabrizian P, Baran PK, Smith WR, Meentemeyer RK. Exploring perceived restoration potential of urban green enclosure through immersive virtual environments. J Environ Psychol [Internet]. 2018;55:99–109. Available from: https://doi.org/10.1016/j.jenvp.2018.01.001
8. Zhang F, Zhou B, Liu L, Liu Y, Fung HH, Lin H, et al. Measuring human perceptions of a large-scale urban region using machine learning. Landsc Urban Plan [Internet]. 2018;180(October 2017):148–60. Available from: https://doi.org/10.1016/j.landurbplan.2018.08.020
9. Salesses P, Schechtner K, Hidalgo CA. The Collaborative Image of The City: Mapping the Inequality of Urban Perception. PLoS One. 2013;8(7).
10. Kyriakou K, Resch B. Spatial Analysis of Moments of Stress Derived from Wearable Sensor Data. Adv Cartogr GIScience ICA. 2019;2(November):1–8.
11. Osborne T, Jones PI. Biosensing and geography: A mixed methods approach. Appl Geogr [Internet]. 2017;87:160–9. Available from: http://dx.doi.org/10.1016/j.apgeog.2017.08.006
12. Aiello LM, Schifanella R, Quercia D, Aletta F. Chatty maps: Constructing sound maps of urban areas from social media data. R Soc Open Sci. 2016;3(3).
13. Alvarsson JJ, Wiens S, Nilsson ME. Stress recovery during exposure to nature sound and environmental noise. Int J Environ Res Public Health. 2010;7(3):1036–46.
14. Hedblom M, Knez I, Ode Sang, Gunnarsson B. Evaluation of natural sounds in urban greenery: Potential impact for urban nature preservation. R Soc Open Sci. 2017;4(2).
15. Hedblom M, Gunnarsson B, Iravani B, Knez I, Schaefer M, Thorsson P, et al. Reduction of physiological stress by urban green space in a multisensory virtual experiment. Sci Rep. 2019;9(1):1–11.
16. Russell JA, Pratt G. A description of the affective quality attributed to environments. J Pers Soc Psychol. 1980;38(2):311–22.
17. Lee EH. Review of the psychometric evidence of the perceived stress scale. Asian Nurs Res (Korean Soc Nurs Sci) [Internet]. 2012;6(4):121–7. Available from: http://dx.doi.org/10.1016/j.anr.2012.08.004
18. Helbich M, Yao Y, Liu Y, Zhang J, Liu P, Wang R. Using deep learning to examine

street view green and blue spaces and their associations with geriatric depression in Beijing, China. Environ Int [Internet]. 2019;126(October 2018):107–17. Available from: https://doi.org/10.1016/j.envint.2019.02.013
19. Queralt A, García JM, Pérez MT, Cerin E, Barnett A, Timperio A, et al. Reliability of streetscape audits comparing on - street and online observations: MAPS - Global in 5 countries. Int J Health Geogr [Internet]. 2021;1–11. Available from: https://doi.org/10.1186/s12942-021-00261-5
20. Scopelliti M, Carrus G, Bonaiuto M. Is it really nature that restores people? A comparison with historical sites with high restorative potential. Front Psychol. 2019;9(JAN):1–12.
21. Barkow JH, Cosmides L, Tooby J, editors. The adapted mind: Evolutionary psychology and the generation of culture. The adapted mind: Evolutionary psychology and the generation of culture. New York, NY, US: Oxford University Press; 1992. xii, 666–xii, 666.
22. Joye Y, van den Berg A. Is love for green in our genes? A critical analysis of evolutionary assumptions in restorative environments research. Urban For Urban Green [Internet]. 2011;10(4):261–8. Available from: http://dx.doi.org/10.1016/j.ufug.2011.07.004
23. Dosen AS, Ostwald MJ. Evidence for prospect-refuge theory: a meta-analysis of the findings of environmental preference research. City, Territ Archit. 2016;3(1):1–14.
24. Stemps III AE. Some Findigns on Prospect and Refuge Theory: II. Percept Mot Skills. 2008;107(1):141–58.
25. Mealey L, Theis P. The relationship between mood and preferences among natural landscapes: An evolutionary perspective. Ethol Sociobiol. 1995;16(3):247–56.
26. Moura JMB, Júnior WSF, Silva TC, Albuquerque UP. The influence of the evolutionary past on the mind: An analysis of the preference for landscapes in the human species. Front Psychol. 2018;9(DEC):1–13.
27. MacKerron G, Mourato S. Happiness is greater in natural environments. Glob Environ Chang [Internet]. 2013;23(5):992–1000. Available from: http://dx.doi.org/10.1016/j.gloenvcha.2013.03.010
28. Wood E, Harsant A, Dallimer M, de Chavez AC, McEachan RRC, Hassall C. Not all green space is created equal: Biodiversity predicts psychological restorative benefits from urban green space. Front Psychol. 2018;9(NOV):1–13.
29. Maas J, Spreeuwenberg P, Van Winsum-Westra M, Verheij RA, de Vries S, Groenewegen PP. Is green space in the living environment associated with people's feelings of social safety? Environ Plan A. 2009;41(7):1763–77.
30. Gatersleben B, Andrews M. When walking in nature is not restorative-The role of prospect and refuge. Heal Place [Internet]. 2013;20:91–101. Available from: http://dx.doi.org/10.1016/j.healthplace.2013.01.001
31. Wilson JS, Kelly CM, Schootman M, Baker EA, Banerjee A, Clennin M, et al. Assessing the built environment using omnidirectional imagery. Am J Prev Med [Internet]. 2012;42(2):193–9. Available from: http://dx.doi.org/10.1016/j.amepre.2011.09.029
32. Rundle AG, Bader MDM, Richards CA, Neckerman KM, Teitler JO. Using google street view to audit neighborhood environments. Am J Prev Med [Internet]. 2011;40(1):94–100. Available from: http://dx.doi.org/10.1016/j.amepre.2010.09.034
33. Badland HM, Opit S, Witten K, Kearns RA, Mavoa S. Can virtual streetscape audits reliably replace physical streetscape audits? J Urban Heal. 2010;87(6):1007–16.
34. BicycleDutch. Reconstruction of a former main street in Utrecht [Internet]. 2020. Available from: https://bicycledutch.wordpress.com/2020/11/04/reconstruction-of-a-

former-main-street-in-utrecht/
35. Wooller J-J, Barton J, Gladwell VF, Micklewright D. Occlusion of sight, sound and smell during Green Exercise influences mood, perceived exertion and heart rate. Int J Environ Health Res [Internet]. 2016 May 3;26(3):267–80. Available from: https://doi.org/10.1080/09603123.2015.1109068
36. Annerstedt M, Jönsson P, Wallergård M, Johansson G, Karlson B, Grahn P, et al. Inducing physiological stress recovery with sounds of nature in a virtual reality forest - Results from a pilot study. Physiol Behav [Internet]. 2013;118(2013):240–50. Available from: http://dx.doi.org/10.1016/j.physbeh.2013.05.023
37. Yatmo YA. Perception of street vendors as 'out of place' urban elements at day time and night time. J Environ Psychol [Internet]. 2009;29(4):467–76. Available from: http://dx.doi.org/10.1016/j.jenvp.2009.08.001
38. Rohrmann B, Bishop I. Subjective responses to computer simulations of urban environments. J Environ Psychol. 2002;22(4):319–31.
39. Williams DR. Making sense of 'place': Reflections on pluralism and positionality in place research. Landsc Urban Plan [Internet]. 2014;131:74–82. Available from: http://dx.doi.org/10.1016/j.landurbplan.2014.08.002
40. Williams DR. "Beyond the commodity metaphor," revisited: Some methodological reflections on place attachment research. Place Attach Adv Theory, Methods Appl. 2013;89–99.
41. Leong YC, Hughes BL, Wang Y, Zaki J. Neurocomputational mechanisms underlying motivated seeing. Nat Hum Behav [Internet]. 2019;3(9):962–73. Available from: http://dx.doi.org/10.1038/s41562-019-0637-z
42. Deroy O, Spence C, Noppeney U. Metacognition in Multisensory Perception. Trends Cogn Sci [Internet]. 2016;20(10):736–47. Available from: http://dx.doi.org/10.1016/j.tics.2016.08.006
43. Bielik M, Schneider S, Kuliga S, Valášek M, Donath D. Investigating the effect of urban form on the environmental appraisal of streetscapes. SSS 2015 - 10th Int Sp Syntax Symp. 2015;(July).
44. Joye Y. Architectural Lessons From Environmental Psychology: The Case of Biophilic Architecture. Rev Gen Psychol. 2007;11(4):305–28.
45. Valtchanov D, Ellard CG. Cognitive and affective responses to natural scenes: Effects of low level visual properties on preference, cognitive load and eye-movements. J Environ Psychol [Internet]. 2015;43:184–95. Available from: http://dx.doi.org/10.1016/j.jenvp.2015.07.001
46. Nilsson K, Sangster M, Gallis C, Hartig T, Vries S, Seeland K, et al. Forests, trees and human health. Forests, Trees and Human Health. 2011. 1–427 p.