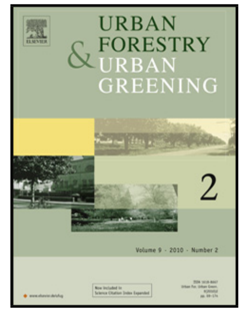
Accepted Manuscript



Title: How eye-catching are natural features when walking through a park? Eye-tracking responses to videos of walks

Authors: Marco Amati, Jodi Sita, Ebadat Parmehr, Chris Mccarthy

PII: S1618-8667(17)30354-0
DOI: https://doi.org/10.1016/j.ufug.2017.12.013
Reference: UFUG 26041

To appear in:

Received date: 5-6-2017
Revised date: 11-12-2017
Accepted date: 19-12-2017

How eye-catching are natural features when walking through a park? Eye-tracking responses to videos of walks

Marco AMATI
Centre for Urban Research,
RMIT University, Swanston Street, Melbourne, VIC 3000 Australia
Tel. +61 3 9925 9887
Marco.Amati@rmit.edu.au (corresponding author)

Jodi SITA
School of Science
Australian Catholic University,115 Victoria Pde, Fitzroy, Melbourne, Australia 3065
Tel. +61 3 9230 8675
jodi.sita@acu.edu.au

Ebadat PARMEHR
Babol Noshirvani University of Technology, Shariati Ave. Babol, Mazandaran, Iran, +98 9125262852, egparmehr@gmail.com

Chris MCCARTHY
School of Software Systems and Electrical Engineering,
Swinburne University of Technology, John Street, Hawthorn, VIC 3122 Australia
Tel. +61 3 9214 8927
cdmccarthy@swin.edu.au

Since the 1960s researchers have developed a range of techniques for evaluating landscape preference. In parallel with this trend, eye-tracking technology has become cheaper, more mobile and more accurate, heralding a new era of big data capture and analysis for landscape preference. In this project our objective was to capitalise on the increasing mobility, sophistication and cheapness of eye-tracking technology to examine its utility in analysing landscape preference. In the following we describe how we eye-tracked 35 participants as they viewed walks through two different parks in the urban center of Melbourne, Australia. We show how participants dwelt on trees and bushes more than other objects. When we compared this to the time and space that objects occupy, participants overwhelmingly dwelt on artificial objects such as lamp-posts, distant buildings and benches. Overall we provide an exploration and method for analysing eye-tracking data in parks by normalising the dwell time by the content, providing a robust means of comparing different dynamic stimuli such as videos.

Keywords: landscape appreciation, park preference, eye-tracking, machine learning

Introduction: the urban age meets the age of biology

As global urban populations continue to increase, residents will face a number of challenges that include the need for restorative environments and an enhanced connection with nature. The quality of these connections are going to be moderated by what the Organisation for Economic Cooperation and Development has heralded as the ‘age of biology’ in which knowledge of the life sciences is going to dramatically improve lives (OECD, 2001). For citizens of the urban age, as Sussman and Ward (2016) argue, the life sciences will play a key role in tackling city challenges.

Already decades of life sciences research exist to shape the design of green spaces and parks, particularly for landscape researchers. The connection between the extent to which participants look at nature and their well-being has been central to empirical psychological studies stretching back to the 1950s (Aoki, 1999). Refined theories predict how looking at nature will relieve stress and affect behaviour. For example, Kaplan's (1995) Attention Restoration Theory (ART) predicts that natural environments will invite: an attentional bias; a relatively effortless mode of attention; and an aesthetically pleasurable feeling. These different dimensions have been validated empirically (Joye et al., 2013; Berto et al., 2008).

Eye-tracking provides a means to evaluate how and why nature is appreciated, but unpacking vision is highly complex. Buswell (1935), demonstrated that viewers unconsciously adjust their eye-movements to the demands of attention during a visual experience. They will fixate objects not just randomly around a scene, but instead cluster around informative or areas of interest (Buswell, 1935). Since that early study, the definition and significance of eye movements such as saccades and fixations, have sustained continued research interest. Detailed explanations of eye-movements can be found in Holmqvist et al. (2011) with a number of key references providing foundations, techniques and applications (Duchowski, 2017; Holmqvist et al., 2011; Liversedge et al, 2011). At the same time, mid-Twentieth Century, theories of knowledge proposed that the brain's introspection is separate from perception (Barsalou, 2008). However, theories of grounded cognition have helped research move away from these elegant formalisms. Currently vision and the eye-movements that support this are seen as an integrated whole, with the eye movements embodied in the activity of the brain. For example, eye motions are stimulated by fictive descriptions that indicate difficult ('the desert is hilly') compared to descriptions that are easy ('the desert is flat') when viewing the same picture (Richardson and Matlock, 2007). Despite a range of technological, theoretical and empirical challenges, since the 1950s eye tracking has become widely used to analyse behaviour and attention in areas such as, marketing, user experience, psychology, neuroscience, human-computer interaction and film (Horsley, et al, 2014; Bojko, 2013; Burch et al, 2017; Dwyer et al, 2018).

The challenge of deriving eye-tracking data from dynamic stimuli and naturalistic settings

This paper aims to contribute to this development, while exploring a method to evaluate landscape theories using as naturalistic a setting as possible. A key consideration is the choice of static or dynamic stimuli. Static images capture many of the important properties in a natural scene, such as the overall visual complexity, colour and levels of lighting. When compared to dynamic images they allow for better control over constantly changing points of view and changing light conditions. In addition, dynamic stimuli and moving scenes generate much larger data-sets presenting challenges in data processing and analysis (Blascheck et al, 2017; Browning, et al 2016). Overwhelmingly, eye-tracking studies have tended to use static stimuli to examine gaze and attentional relationships in a controlled way. A recent review of eye-tracking data analysis techniques considered published studies from 2010 to the 1950s, revealed only eight previous studies out of a potential 90 that used dynamic stimuli (Blascheck et al., 2014). Yet, as Dorr et al (2010) argue, static stimuli do not generate natural viewing behaviour.

A further consideration is whether the study has been framed around asking participants to perform a particular task or allowing participants to freely view a stimulus. Nordh et al. (2013) note that many studies in landscape and environmental psychology have used the latter. Yet in a research environment participants are likely to view the stimulus in a mixture of free viewing and task oriented modes. For example, even when instructed to freely view a picture of a room to evaluate say, the effect of colour, participants' eye-tracking data may be affected by extraneous considerations such as whether they notice if the room is tidy (Geisler and Cormak, 2011). This bias and noise may overwhelm a signal when the stimulus is dynamic (Mital et al. 2011). In an outdoor setting, the situation becomes more complex still as the participant has the freedom to change their field of view. Significant challenges remain in using more complex and naturalistic settings to understand eye-tracking behaviour. Nonetheless, if the potential of this data to explain landscape theories is to be realised, a necessary step is to be able to control for variability in the field of view between participants and stimuli.

In the current study, we showed participants high-definition video footage of walks through two urban parklands in the City of Melbourne, Australia. There are many eye-tracking measures which might be used in a study such this, these include: fixation frequency to return to an area of interest and average fixation duration (Jacob & Karn, 2003). A fixation can be defined as a stable position relative to an object for a certain amount of time. It is an attempt to keep light steady on the macula of the retina. Some measures such as fixation duration and fixation frequency are strongly correlated. We decided to use the fixation time for an area of interest because this has been shown to be positively correlated with participants' assessment of the potential of a park to be restorative (Nordh et al. 2013). Fixation time for a given object can be summed to produce dwell time. Our aim was to capture the dwell times of participants while controlling for the size and time of exposure of different elements, such as trees and grass to be able to compare the two parks. Our specific questions were:

1. How different is the dwell time for different areas of interest (AOI), such as trees and rocks, during the course of a single park walk?
2. How does the dwell time on an AOI vary between parks?
3. How do these two comparisons (1. and 2.) vary when we consider the amount of exposure of the AOI to the participants in the different videos?
4. What are the links between the participants' opinion of the parks and their dwell time for the dynamic stimuli?

A professional film production company was hired to video record similarly timed walks through two different parks between March to May 2015. We asked 35 participants to view videos of the walks while imagining themselves in need of relaxation. We recorded the participants' eye-tracking data and used an Area of Interest based approach to analyse the data. Participants were also asked for their opinion and rating of the different walks. Ethics approval was gained by ACU Human Ethics Committee (2015-36E).

Method:

Phase 1: selecting parks for video recording

In the first phase the researchers in consultation with the City of Melbourne (CoM) selected a list of potential parks for analysis, these were narrowed to Royal Park and Fitzroy Gardens (Figure 1). These parks represent a very diverse set of potential scenes, from the formal, English pastoral park setting of Fitzroy Gard ens to the more wilder setting of Royal Park, dominated by local native planting.

In consultation with the CoM, filming routes and a sitting scene within the parks were selected. So as to maintain comparability between the parks, films were produced on the following basis:

- A professional filming company (https://www.bakewood.com.au) was employed to produce digital films for the project. A high quality digital video camera, (Panasonic GH4 with an Olympus 12mm lens, standard 35mm equivalent of 24mm) was attached to a Nebula Lite 4000 gimbal and used to create each of the simulated 'walking' or simulated 'sitting' videos with smooth fluid movements in in 4k resolution at 25 fps. The final videos delivered were 1920 x 1080 HD.
- Films were shot in one take, to a standard length with precise predefined routes. Each simulated walk film was 3 mins 50 seconds long and each simulated sitting film was 1 minute long for both parks. Overall the aim was to ensure that the films were similar and did not introduce novel gaze data due to different walk trajectories or camera movement. The sitting and walking scenes were input as separate files so as to test them differently in the Tobii software.
- As much as possible the scenes were filmed without the intrusion of other people, vehicles and excessive noise (e.g. construction noise); since it was assumed that this would affect the eye-tracking results. In practice, some of this intrusion did occur so the relevant video frames were excluded from the analysis. The ambient sounds in both videos were similar and comprised the sound of light winds, ambient traffic noise and birds.

Phase 2: selecting participants for the study

Following the receipt of ethics approval the team recruited 37 participants for the study, however for one of these participants' eye-tracking data was not valid and another failed to come to the appointment. Twenty of the participants were recruited using a professional research recruitment company to ensure a spread of different genders, ages and occupations among the sample. A further 17 were recruited on site from among the student and staff body. Participants had a median age of 26-35, with 46% and 56% comprising females and males respectively (Table 1). The participants received a small gift for participating in the study (Australian $50).

Participants were surveyed about their age, gender and frequency of park use before being asked to imagine themselves in need of restoration. The exact wording on the screen was as follows:

'Imagine that it is midday and you are walking alone in Melbourne. You are mentally tired from intense concentration at work and are looking for somewhere to go for a walk, sit down and rest for a little while, before going back to work'.

Participants were seated comfortably in front of the computer screen at a distance of 60-65 cm to view the films and paired speakers were set up to play the audio from the films. A Tobii x120 remote eye tracking device was used to record eye-movements at 120 HZ as the participants watched the footage on a 16 inch flat screen PC Dell monitor in high definition. Tobii Studio version (2.1.14) was used to set up the study, calibrate and record the participants' eye-movements. Participants were calibrated to the eye tracker using an automated 5 point calibration procedure. The ordering of the two park films were randomised using Excel (Microsoft Corporation) so as to ensure that no bias would be generated by one park film being consistently used first or last. Although with only two parks, it would be possible to alternate the order of the parks instead of randomising them, this study was part of a larger project utilising other film based stimuli. Recording the data from each participant took about 20 min per participant. This included testing, calibration watching the films and answering the questions.

Phase 3: triangulating the eye-tracking data with other information

At the end of each film participants were asked to rate each of the parks on a scale of 1 to 10. Firstly for whether they felt they would be able to rest and recover in that environment (1 not very much- 10 very much) and secondly how much did they like the park (1 not very- 10 very much). They were also asked whether they would use the park to relax or relieve stress (Y/N) and whether they recognised the park they were looking at (Y/N). The recognition was to control for the potential impact of a prior association with the park. In addition, short answer responses were recorded using a voice recorder and then transcribed to explain these impressions.

Phase 4: quantifying the content of the videos

According to a review, past research has used two principle methods of analysing eye-tracking data (Adrienko et al. 2012). Point-based analysis examines the spatial or temporal distribution of the eye movement data. Area of Interest (AOI) based analysis involves identifying areas in the stimulus that are of interest and generating data that are specific to those areas. A specific variation on AOI analysis is that of Nordh et al. (2013), who used relative AOI: the amount of time spent dwelling on one AOI in photos of a scene divided by the total dwell time across all AOIs. These authors also used the surface area of the AOIs in each image to express the dwell time relative to the area of the AOI. This approach can allow eye-tracking data from scenes with varying AOIs to be compared, since by controlling the eye-tracking data for a given AOI by the amount of content, the opportunity for the participant to examine an AOI remains the same.

We aimed to analyse our dynamic stimuli of Royal Park and Fitzroy Gardens video walks in a similar way to Nordh et al. (2013) by subjecting it to a detailed content analysis along the full length of the video. The video of the walks in these parks were broadly similar because they contained upper and lower stories, both contained rock features along the borders of the walking paths, and also similar sequences of events, i.e. a flight of steps early on in the walk and then a winding path that opened onto more open scenery over the park itself. In addition, the videos were filmed from the same height (1.5 m) and there were no landscape elements that were unique to either park (e.g. one park did not have a view of water and the other not).

At the same time, the park walk videos are significantly different from one another because of the age and type of planting. The Fitzroy Gardens video contains a great deal of lush green vegetation and is darker, with less sky. The vegetation is planted close to the path and at times creates a tunnel effect. Where the view opens up it is of lawns and dispersed deciduous trees. The Royal Park walk contains an identifiably South-East Australian mixture of native vegetation suited to a drier climate. It is more consistently open, with more sky. The grass is longer and the colour palette varies between yellows and light greens. Given the parks' similarities and differences they enable a contrast between a landscape dominated by exotic European plantings and one dominated by Australian native flora.

To analyse the content in the video we employed an automatic multi-class image segmentation technique (Gould, 2012). This technique was chosen on the basis that it provides robust pixel labelling in a known video sequence with low training requirements. The image segmentation method used in this study is representative of a class of techniques that employ Markov Random Field analysis to perform semantic labelling in real world pixel data. These approaches incorporate both local evidence for a particular object class, while simultaneously smoothing out noise in local labelling using global constraints. The variant of the method applied here has been recently applied in a low vision assistive vision context, in which several classes of objects on natural scenes were labelled (Horne, et al. 2016).

First, we selected approximately 30 image frames from both park videos to train the system. Training images were generated using a given still from the video and were painted manually in GIMP (GNU Image Manipulation Program) software for 1% of all frames in the video file (Figure 2). Given that the training frames were directly drawn from the same sequence of images as were to be processed, and were sampled at points in the sequence that best represented the diversity of appearance of each AOI, 30 frames was estimated to provide sufficient accuracy. This estimate was validated via manual inspection of randomly selected output frames across the sequence (approximately every 50 frames). The use of fewer frames to train would be expected to degrade the accuracy of the classifier for less common object classes such as artificial objects, which only appear intermittently. By contrast, AOIs such as the path require very few frames to achieve a reasonable statistical model for classification.

The pre-training of the machine learning-based image segmentation algorithm was used to identify and label pixels belonging to specific Areas of Interest: upper story vegetation; paths; rocks; artificial objects; lower story vegetation and sky (Table 2). From this, machine learning was used to train a pixel classifier, which was then run over every frame of the original video sequences, generating labeled AOIs for each frame.

Figure 3 shows one frame from each park with the corresponding pixel labelling as a demonstration of the accuracy of the classification.

Phase 5 Combining the eye-tracking data with the quantitative data on the content of the video

In the final phase, the raw eye tracking data for each subject was imported into software (Tobii Studio 3.2.1) for data filtering and fixation determination. A filter within in the software, the I-VT filter (classifier: 30°/s; Velocity calculator window length: 20 ms), was used to filter the data and classify the fixations.

The eye-tracking data were precisely matched with the video content using a reference cross in the top of the screen and some distinct objects within the video. The relative AOIs were calculated in 5 second blocks using the classified pixels. The number and length of fixations on the different objects were calculated for the same blocks. The amount of object there is in a video, either in 5 second block or as a whole, can then be compared to the dwell time that the object attracts.

Results

Analysing the contents of the park videos

The percentage of each AOI as content and dwell time is shown with a different colour for each 5 second window across the length of each video (Figure 4). The different categories were upper-story vegetation (dark green), path (light brown), rocks (dark brown), artificial (red), lower-story vegetation (light green) and sky (blue). As can be seen in Figure 4 (a) the Royal Park video has more sky visible than the Fitzroy Gardens video. Conversely, the Fitzroy Gardens video contains a large amount of lower story (turf) at the start before the path enters a part of the park which is densely foliated (Figure 4 (b)). By comparing the bar charts of the video content (Figure 4 (a) and (b)) with the dwell time for each AOI (Figure 4 (c) and (d)), it is clear that artificial objects, for example, attract proportionally more visual attention compared to the area they occupy in the video. Similarly, sky appears consistently along the length of the Royal Park video, however is looked at in only a few cases (Figure 4 (a) and (c)).

Analysing the extent to which participants viewed different elements in the walks for all participants

Total Dwell times

We extended the analysis for participant A to all 35 participants in the next stage of analysis. Figure 5 shows shows a box plot of all the participants' dwell time for the different AOIs in both parks. The plot shows that vegetation AOIs in the form of bushes or trees are dwelt on the most in Fitzroy Gardens and Royal Park videos respectively. On the other hand, a relatively low amount of dwell time is spent on artificial and sky AOIs.

The data are not normally distributed so a Kruskal-Wallis test and a Dunn's post-hoc test was used to show that there exists a statistically significant difference between the AOIs in each park. The tests revealed that there was at least one dwell time on an AOI that was different to the others (for Fitzroy Gardens AOIs X^2 = 192.289 p .000 Dunn's post hoc p = .000 for Royal Park AOIs X^2 = 180.428 p .000 Dunn's post hoc p = .000).

A Mann-Whitney U test was then employed to see whether there are significant differences for each category between the parks (Table 3). The table shows that, statistically, more time was spent looking at artificial objects, rocks, sky and trees in Royal Park compared to Fitzroy Gardens. On the other hand participants dwelt longer on bushes in Fitzroy Gardens than Royal Park. There is clearly more sky visible in Royal Park than Fitzroy Gardens and more bushes visible in Fitzroy Gardens than Royal Park, which partially explains this result (Figure

4 (a) and (b)). Neither path in each park attracts more dwell time than the other, demonstrating the similar role that this content plays in guiding the participants' eyes.

Total Dwell times divided by the content

In the next stage of the analysis we divided the dwell times by the area of the AOIs. Since the content of the videos differ considerably, the differences in the dwell times can be explained by the size of the area or opportunity to view certain objects. For example, more sky is visible in Royal Park so the dwell times are higher on sky for Royal Park when compared to Fitzroy Gardens. By dividing by the content in the video we were able to normalise the dwell time across both park walks for this effect.

Figure 6 shows a very different spread of results compared to Figure 5. In both parks the dwell time/content on the artificial AOI is considerably higher than other AOIs, particularly in Royal Park. In other words, objects that occupy a relatively small area of the field of view during the course of the video, such as light poles and benches are dwelt on to a far greater extent compared to their size, when compared with objects that dominate the field of view such as trees.

The Kruskal-Wallis and post-hoc Dunn's tests again revealed that there was an overall difference between the dwell times on the AOI (for Fitzroy Gardens AOIs X^2 = 126.910 *p* .000 Dunn's post hoc *p* = .000 for Royal Park AOIs X^2 = 138.848 *p* .000 Dunn's post hoc *p* = .000).

Similar to Table 3, Table 4 shows the result of the Mann-Whitney U test to compare the participants' views of both parks. Table 3 shows some differences that occur by dividing the dwell times by the content. Firstly, the dwell time on the path is even more similar among both parks by dividing the dwell time by the content than using the dwell time alone (M-W test *p* .986 when dividing dwell time/content compared to M-W test *p* .703 for dwell time alone, Table 3, 4). Secondly, by dividing the dwell time by the content there is no statistical difference between the view of the rocks (*p* .394). This can be explained because many of the rocks line the path and therefore guide the eye in both ways across both parks. All the other dwell time AOI categories are significantly different comparing the parks (Table 3, 4). An exception to this is sky in Fitzroy Gardens which is dwelt on less overall compared to sky in Royal Park. However, in comparison to the amount of sky in both parks the sky in Fitzroy Gardens is dwelt on relatively more than Royal Park.

Table 4: Comparison for Dwell time of view categories divided by the content between two parks (Mann-Whitney U-Test) (FG: Fitzroy garden; RP: Royal park)

Dwelling on a stair climbing sequence

While the path is looked at to a similar extent in both parks we decided to examine whether this varied with the stair climbing sequence (Figure 7). As Land & Tatler (2009: 103) show when walking up stairs, it is often only the first few stairs that are looked at. In our study for Royal Park the stair climbing sequence began after 45 seconds, and after 75 seconds in the case of Fitzroy Gardens, both sequences last approximately the same amount of time. As Figure 8 shows, even though our study is desktop based, the average dwell time on the path, which also includes stairs, peaks in the initial stage of the sequence across all 35 participants and then there is a drop off along the course of the sequence.

Analysing the participants' reactions to the park

Participant responses to the survey questions are summarized (Table 5). Overall it is clear that the simulated walk through Fitzroy Gardens was preferred over that of Royal Park. The results also reveal a significant variety of reactions to the walks. The lowest rating for Fitzroy Gardens was associated with a view of the walk as being dark and oppressive. On the other hand, Royal Park's dry and open native plantings were also seen to be oppressive because the park seemed lonely and unkempt. Clearly, the parks' contrasting vegetation colours had a significant role to play in affecting the participant reactions. The browns and yellows of Royal Park were less favoured over the rich deep greens of Fitzroy Gardens. The dark rich vegetation of Fitzroy Gardens led some participants to feel more secluded. At the same time, while seclusion led to feelings of oppression in some participants others saw this as an advantage.

In the final stage of the study we examined the links between the participants' opinion of the park and their dwell time on different objects. Table 6 shows the participant ratings of each park and their correlation with the time subjects spent dwelling on each AOI. In Fitzroy Gardens, where subjects rated the park highly, there is a positive and significant correlation (.486) for dwelling time on Trees. For the Royal Park gardens, there was a significant positive correlation between opinion of the park and the dwell time on sky (.337) as well as a significant negative correlation with dwell time towards the path (-.407).

Discussion

Eye-tracking data represents a potentially powerful way of eliciting responses from people about their environments. In an increasingly urban world, the quality and quantity of people's 'green dosage' matters physiologically and psychologically (Thompson et al. 2012). Looking at nature represents an important means of delivering that dosage.

Previous comparable studies have used the potential AOIs that participants can view and compared this with actual views in naturalistic settings. For example Gidlöf et al. (2012) examined the proportion of advertisements that teenagers look at and are aware of, compared to the actual advertising on websites. Gidlöf et al. (2013) also examined decision-making by purchasers in a naturalistic mode by using a supermarket as a stimulus. However, we know of no study that has attempted this for a walk in a park. This project and its unique data analysis method has allowed us to render feasible the collection and comparison of moving image eye-tracking data from an outdoor setting. Normalizing the data according to what can potentially be seen by the participants, has allowed us to probe comparative viewing behaviour in new ways. Findings gained though this analysis indicate a confirmation of prior work on visual saliency and landscape design in a moving paradigm as we acknowledge below. We further suggest how this method can be of value in other research contexts.

Confirming visual saliency and considering way-finding

In a seminal work Itti et al. (1998) showed how the eye's attraction to different objects based on edges, contrast and colour fits a model in a free-viewing situation. Since then a large body of work has further refined these findings (e.g. Carmi and Itti, 2006; Torrabala et al, 2006; Anderson et al, 2015). These works collectively help to explain some of the emphasis on artificial objects in the videos, such as light poles, which have sharper edges and are a contrasting colour to the background. In previous study using landscape photographs Dupont et al. (2016) use visual saliency to argue that peri-urban scenes attract the eye more because they contain greater light and colour contrast between artificial objects, such as houses and natural objects such as fields. The darker colours of our artificial objects are especially prominent against the lighter background vegetation of the Royal Park film, explaining why they are dwelt on more.

Further work on scene viewing shows that while the factors predicted by visual saliency models are important and often correlated with fixation behavior, it is only one component at play (Einhauser et al, 2007, Henderson et al, 2007, Tatler, 2007) and that high-level cognitive factors are also important in guiding gaze behavior in natural scenes (Henderson, 2003; Peters et al, 2005; Siebold et al., 2011). Natural scenes are rich in detail and meaning and have complex spatial arrangements. In our study while watching the guided park walk, participants were considering whether they recognized the park. Recognition for urban dwellers comes from distinguishing artificial objects and their placement, especially when these artificial objects include signs or large objects outside the park. Thus attention to these objects may be because the eye was first drawn by their salience, then dwelt on longer due to processing what the object was and whether they recognized it or could place it. Equally, we acknowledge that dwell time may be a marker for many cognitive processes, not only categorization and recognition. Further research is required to confirm how eye-tracking data can be used to evaluate recognition in a natural scene such as this.

Other aspects of the videos that may have influenced the participants' eye-tracking behaviour may be due to the act of walking or climbing the stairs in the videos. In Fitzroy Gardens, the darker path between the thickly vegetated garden beds may have provoked anticipation and mystery causing the respondents to examine the edges of the path, defined by rocks, as the scene is revealed. This viewing behaviour is similar to a navigation task such as driving a car, where the drivers tend to look ahead on the path and into the corners of bends in the road (Land & Tatler, 2009: 119-121). In Royal Park artificial objects may have provided a distant reference point aiding the process of recognition. Our study was a younger cohort with a large proportion of 26-35 year olds and did not include anyone over 55. The pattern of fixations for stair climbing may have been different with an older (median age >70) cohort (Zietz and Hollands, 2009).

Comparing the eye-tracking data to the rating of the park videos

While we found some significant correlations between AOIs in the parks the interpretation of these is rendered difficult because of the length of the videos and the experimental design. The more participants looked at trees in Fitzroy Gardens the more likely they were to positively evaluate the park. Given that participants favoured Fitzroy Gardens more than Royal Park it may be that the shape, age and foliage of the European trees in Fitzroy Gardens contributed to this opinion. This is confirmed in a number of the short answer responses. A positive attribute of the Royal Park gardens walk was its openness and participants who viewed the sky more were more likely to rate the park positively. However, openness as a property is to do with the depth of field as well as the sky and further research is required to understand the reasons behind these opinions. Finally, the path materials are both different in both parks. In Fitzroy Gardens it is asphalt but a gravel and sand mix in Royal Park. The more participants looked at the Royal Park path, the less they liked it and may have contributed to their view of it as being unkempt.

Some of the trends that are found in the comments, for example the closeness of the vegetation in the case of Fitzroy Gardens can only be loosely connected to the eye-tracking data. For example, for the one negative perception of Fitzroy Gardens as being 'closed in' may explain why in comparison to the amount of sky in both parks, the sky in Fitzroy Gardens is dwelt on relatively more than in Royal Park. Alternatively, the feeling of being 'closed in' may be a product of involuntarily being forced to look at features close to the path.

Clearly, a more refined and focused approach to the experimental design is required to unpack how participants' eye-tracking data varies with a stimulus such as this. The lack of overlap between eye-tracking and self-reported evaluations to the parks is in contrast to Nordh et al. (2013), who found a good correlation between eye-tracking on photos and self-reported evaluations. Dynamic video footage lasting close to four minutes would prompt so many more mixed conscious, conscious action (top-down processing drivers) and unconscious (bottom-up drivers) fixation points than a static image. Teasing out and relating this wealth of data to a self-reported Likert scale at the end of the video is problematic because, for example, participants are likely to remember and be influenced by the final scenes of the video.

One way to explain the eye-tracking in a video of this length would be to replay the eye-tracking data to the participants and ask them to reflect on and remember why they were looking at certain objects. A concurrent method of reflection may also reveal a similar data. However, this method may explain the total dwell time, but not the emphasis or the dwell time divided by the content, as we saw with the artificial objects. This suggests that for a retrospective evaluation of eye-tracking data divided by the content it is important to identify this emphasis *in situ* and then ask the participants why they examined the artificial objects to the extent that they did. A further means to increase the immediacy of the landscape rating is to divide the video into segments and ask participants to rate more often as the scene unfolds (Gandy and Meitner, 2007) or the participant could rotate a dial or move a lever to indicate levels of increasing or decreasing approval.

Away from the video screen and in an even more naturalistic, outdoor setting with the full gamut of sensory stimulation at play it is likely that the participants' answers to questions about the walk and their eye-tracking data would be even further apart. As Kroh and Gimblett (1992) noted, infield and lab contexts of a study can drastically affect landscape rating for a walk through the same landscape and with the same participants. Yet, the extent to which the eye-tracking data is similar in a video stimulus and an outside setting is unknown.

Conclusion: implications for landscape practitioners and future applications

Eye-tracking research can be valuable for studying human and urban interactions. This kind of data can have significant impacts in learning more about the urban environment for safety, aesthetics, advertising, expertise and direction finding, to name a few applications. Furthermore, the cost of eye-tracking equipment has reduced and made it accessible to researchers in a variety of fields and to governments.

Our aim in this study was to examine the extent to which participants viewed elements or AOIs in a video of a walk through a park and question how this data could be analysed. We have shown that the elements themselves are dwelt on differently in a single park and the same elements are dwelt on differently across parks. We have analysed this data both in terms of dwell time and by dividing the dwell time by the total content to show what participants emphasised. We have attempted to interpret the data based on the features in the park and the reactions of the participants to them.

A major challenge we have identified is that eye-movements in a natural setting while walking for close to 4 minutes requires an experimental design that isolates one of three potential research areas. The first, is the relationship between the dwell time, arguably what the eye is seeing and processing and relating this to opinion of scenes. The second, is the emphasis on particular objects over others, that are driven by tasks such as, in our study, orientation. The third is the attenuation of interest on objects, in our case stairs, once a task is understood and the attention will be directed at other objects. Given these directions, to what extent can eye-tracking help landscape practitioners?

Design implications for landscape practitioners

Landscape park designers and practitioners are interested in improving park design and the legibility of these spaces. The study reported here provides an early indication of some of the factors that can affect legibility and direction finding. In the future, we argue that eye-tracking has useful applications in testing well known landscape theories and for teasing out the nature of landscape expertise.

A primary conclusion is the sensitivity of the eye to artificial objects. Even when they occupy a tiny fraction of the scene they can be easily perceived and fixated on by the eye. Therefore, if landscape practitioners are interested in designing a truly immersive and 'natural' scene, eliminating these elements to the greatest amount possible is important. Furthermore, the study showed that the act of walking, implies a narrative of 'direction finding' that unconsciously directs the eye towards way markers. The emphasis on artificial objects in Royal Park and rocks in Fitzroy Gardens, which form the edges of the path, are indicative of the eye responding to the narrative content of the video and seeking a direction as the walk progresses, even when these walks are well known and even when the participants are watching the walks on a screen.

If designers wish to immerse participants in a natural setting providing a guide to the use of natural way markers and camouflaging the intrusion of artificial objects is important. Indigenous Australians, especially the Wurundjeri and other Kulin peoples who traversed and inhabited and continue to inhabit the area we call Melbourne, would have used a plethora of natural cues to orient themselves in the landscape (Orr-Young, 2012). Design traditions in other countries also provide a guide to producing an immersive natural experience. For example, the Japanese concept of *shakkei* or borrowed landscapes.

Additional applications for the methods developed in this study may lie in testing some well known landscape theories. A dominant one is prospect and refuge theory (Appleton, 1975). This predicts that people who associate a scene with a greater fear of crime will do so because of a lack of open spaces or prospects and a dominance of refuges or hiding places for a potential hazard. Defining prospects and refuges using a pixel classifier as this study would enable a relative understanding of the volumes that these spaces occupy in a video and would provide a mechanism to test whether participants' dynamic rating of a fear of crime relates to affects eye-tracking behaviour. A remaining challenge would be to record the fear of crime in a similarly dynamic way. A similar potential application is to test Attention Restoration Theory (Kaplan, 1995), where the quantity of a variety of eye-tracking measures could be related and linked to stress reduction. However, both of these theories are premised on an ambient, free viewing context. Yet, we know that a task oriented research design would provide a clearer way to prove a hypothesis.

Finally, the way in which different cohorts view landscape scenes has been a key interest for several decades (Daniel and Boster, 1976: 37). Eye-tracking could be used to delve more deeply into the differences between cohorts. What do experts for example urban designers or arborists see differently in an urban scene of buildings or trees compared to non-experts? Eye tracking has been used to examine expertise effects in many disciplines and professions (Gegenfurtner, A. *et al*, 2011, Harrison. *et al*, 2016) and in many cases helps to distinguish previously unknown features of skill, understanding or cognitive schema. It has also recently been used to study this question using eye tracking with still photographs (Dupont et al, 2015). An eye-tracking study could help landscape designers better identify what attracts the eye of an expert which could have a pedagogical role.

Acknowledgments:
This work was financially supported by the City of Melbourne. We would like to thank Eamonn Fennessy for his advice and collaborative approach. In addition, A/Prof Adrian Dyer and Dr Qian Sun provided invaluable help and advice for improving the analysis of the data. One of the reviewers worked very carefully through the paper and noted many errors which we hope we have corrected to their satisfaction. All errors remain the authors'.

References:

Anderson, N. C., Ort, E., Kruijne, W., Meeter, M., & Donk, M. (2015). It depends on when you look at it: Salience influences eye movements in natural scene viewing and search early in time. *Journal of Vision*, 15(5), 9.

Andrienko, G., Andrienko, N., Burch, M. & Weiskopf, D. (2012). Visual analytics methodology for eye movement studies. *IEEE Transactions on Visualization and Computer Graphics*, 18, 2889–2898.

Aoki, Y. (1999). Review article: trends in the study of the psychological evaluation of landscape. *Landscape Research*, 24, 85-94.

Appleton, J. (1975). *The experience of landscape*. London: Wiley.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology,* 59, 617–645.

Berto, R., Massaccesi, S. & Pasini, M. (2008). Do eye movements measured across high and low fascination photographs differ? Addressing Kaplan's fascination hypothesis. *Journal of Environmental Psychology*, 28, 185–191.

Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2014). State-of-the-Art of Visualization for Eye Tracking Data. Eurographics Conference on Visualization (EuroVis), no page.

Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2017). Visualization of Eye Tracking Data: A Taxonomy and Survey: Visualization of Eye Tracking Data. Computer Graphics Forum. Vol 00, 0, 1-25 10.1111/cgf.13079.

Bojko, A. (2013). Eye Tracking the User Experience. A Practical Guide to Research. Rosenfeld.

Browning, M., Cooper, S., Cant, R, Sparkes, L., Bogossian, F., Williams, B., O'Meara, P., Ross, L., Munro, G., & Black, B. (2016). The use and limits of eye-tracking in high-fidelity clinical scenarios: A pilot study. *International Emergency Nursing.* Vol.25:43-7. doi: 10.1016/j.ienj.2015.08.002.

Buswell, G,T. (1935). *How people look at pictures.* Chicago, Il; University of Chicago Press.

Burch, M., Chuang, L., Fisher, B., Schmidt,A., & Weiskopf, D. (2017). Eye Tracking and Visualization: Foundations, Techniques, and Applications. ETVIS 2015. Springer International Publishing.

Carmi, R., & Itti, L. (2006). Visual Causes Versus Correlates of Attentional Selection in Dynamic Scenes. Vision Research, 46(26), 4333-4345

Daniel T. C. and Borster, R. S. (1976). Measuring landscape esthetics: the scenic beauty estimation method. USDA Forest Service Research Paper RM-167.

Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision,* 10 (10):28, 1–17

Duchowski, A.T. (2017). Eye Tracking Methodology: Theory and Practice, 3rd ed. Springer International Publishing.

Dupont, L., Antrop, M., & Van Eetvelde, V. (2015). Does landscape related expertise influence the visual perception of landscape photographs? Implications for participatory landscape planning and management. *Landscape and Urban Planning*, 141, 68–77.

Dupont, L., Ooms, K., Antrop, M., & Van Eetvelde, V. (2016). Comparing saliency maps and eye-tracking focus maps: The potential use in visual impact assessment based on landscape photographs. *Landscape and Urban Planning*, 148, 17–26.

Dwyer, T., Perkins, C., Redmond, S., & Sita, J. (2018). Seeing into Screens. Eye Tracking and the Moving Image. Bloomsbury.

Einhäuser W., Mundhenk T. N., Baldi P., Koch C., Itti L. (2007). A bottom-up model of spatial attention predicts human error patterns in rapid scene recognition. *Journal of Vision*. 7, 6.1–6.13

Gandy, R., & Meitner, M. J. (2007). The effects of an advanced traveller information system on scenic beauty ratings and the enjoyment of a recreational drive. *Landscape and Urban Planning*, 82, 85-93.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011) Expertise differences in the comprehension of visualizations: A meta-analysis of eye tracking research in professional domains. *Educational Psychology Review,* 23, 523–552.

Geisler, W.S. & Cormack, L.K. (2011). "Models of overt attention," Oxford handbook on eye movements, eds Liversedge, S.P., Gilchrist, I., Everling, S. (Oxford University Press, Oxford, UK), 439–454.

Gidlöf, K. Holmberg, N. and Sandberg, H. (2012) The use of eye-tracking and retrospective interviews to study teenagers' exposure to online advertising, *Visual Communication,* 11 (3): 329-345.

Gidlöf, K. Wallin, A. Dewhurst, R. and Holmqvist, K. (2013) Using Eye Tracking to Trace a Cognitive Process: Gaze Behaviour During Decision Making in a Natural Environment, Journal of Eye Movement Research 6(1):3, 1-14

Gould, S. (2012). "DARWIN: A Framework for Machine Learning and Computer Vision Research and Development". *Journal of Machine Learning Research (JMLR),* 13, 3533−3537.

Harrison. T. K., Kim, T. E., Kou, A., Shum, C., Mariano. E. R. & Howard, S. K. (2016). Feasibility of eye-tracking technology to quantify expertise in ultrasound-guided regional anesthesia. *Journal of Anesthesia,* 30, 530–533. doi:10.1007/s00540-016-2157-6.

Henderson, J. M. (2003). Human gaze control in real-world scene perception. *Trends in Cognitive Sciences,* 7, 498–504.

Henderson J. M., Brockmole J. R., Castelhano M. S., Mack M. (2007). "Visual saliency does not account for eye-movements during visual search in real-world scenes," in Eye Movement Research: Insights into Mind and Brain, eds Van Gompel R., Fischer M., Murray W., Hills R., editors. (Oxford: Elsevier;), 437–562.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Halszka, J. & van de Weijer, J. (2011). *Eye Tracking: A Comprehensive Guide to Methods and Measures.* Lund: Oxford University Press.

Horne, L., Alvarez, J., McCarthy, C., Salzmann, M. and Barnes, N. (2016). Semantic labeling for prosthetic vision. *Computer Vision and Image Understanding*, *149*, pp.113-125.

Horsley, M., Eliot, M., Knight, B.A. & Reilly, R. (Eds.) (2014). '*Current trends in eye tracking research*' Springer, Switzerland.

Itti, L. Koch, C. and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 20, 1254-1259.

Jacob, R.J., Karn, S.K. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radach, H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research*, Elsevier Science, Amsterdam, pp. 573-605.

Joye, Y., Pals, R., Steg, L., Evans, B. L. (2013). New Methods for Assessing the Fascinating Nature of Nature Experiences. *PLoS ONE*, 8, e65332.

Kaplan, S. (1995). The restorative benefits of nature: Towards an integrative framework. *Journal of Environmental Psychology* 15, 169–182.

Kroh, D. P., & Gimblett, R. H. (1992). Comparing live experience with pictures in articulating landscape preference. *Landscape Research*, 17, 58-69.

Liversedge, S., Gilchrist, I., and Everling, S. (2011). *The Oxford Handbook of Eye Movements*. London, Oxford University Press.

Land, M., & Tatler, B. (2009). *Looking and Acting: Vision and eye movements in Natural Behaviour*. New York, Oxford University Press.

Mital, P.K., Smith, T.J., Hill, R.L., and Henderson, J.M. (2011). Clustering of Gaze During Dynamic Scene Viewing is Predicted by Motion. *Cognitive Computation* 3(1): 5–24. doi:10.1007/s12559-010-9074-z

Nordh, H., Hagerhall, C.M., & Holmqvist, K., (2013). Tracking restorative components. Patterns in eye movements as a consequence of a restorative rating task. *Landscape Research*, 38, 101–116.

Organisation for Economic Cooperation and Development (OECD) (2001) *The Application of Biotechnology to Industrial Sustainability*. Paris: OECD. Available from https://www.oecd.org/sti/biotech/1947629.pdf viewed 25 October 2016.

Orr-Young, J. (2012). *Place and praxis: valuing Australian Indigenous place in landscape architecture practice*, Doctor of Philosophy (PhD), Architecture and Design, RMIT University.

Peters RJ, Iyer A, Itti L, Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Research* 45: 2397–2416.

Richardson, D. and Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition*, 102, 1, 129–138.

Siebold A., van Zoest W., Donk M. (2011). Oculomotor evidence for top-down control following the initial saccade. *PLoS One*, 6 (9), doi:10.1371/journal.pone.0023552.

Sussman, A., & Ward, J. M. (2016). Planning for the subconscious. *Planning: The Magazine of the American Planning Association,* June, 31-35.

Tatler B. W. (2007). The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal Vision.* 7 (4), 1–17.

Torrabala, A., Oliva, A., Castelhano, M.S., and Henderson, J.M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review,* 113, 766-786.

Thompson, C. W., Roe, J., Aspinall, P., Mitchell, R., Clow, A., & Miller, D. (2012). More green space is linked to less stress in deprived communities: Evidence from salivary cortisol patterns. *Landscape and Urban Planning,* 105, 221-229.

Zietz, D. and Hollands, M. (2009). Gaze Behavior of Young and Older Adults During Stair Walking. *Journal of Motor Behavior,* 41 (4), 357-365.

Figure 1: Map of the parks' locations

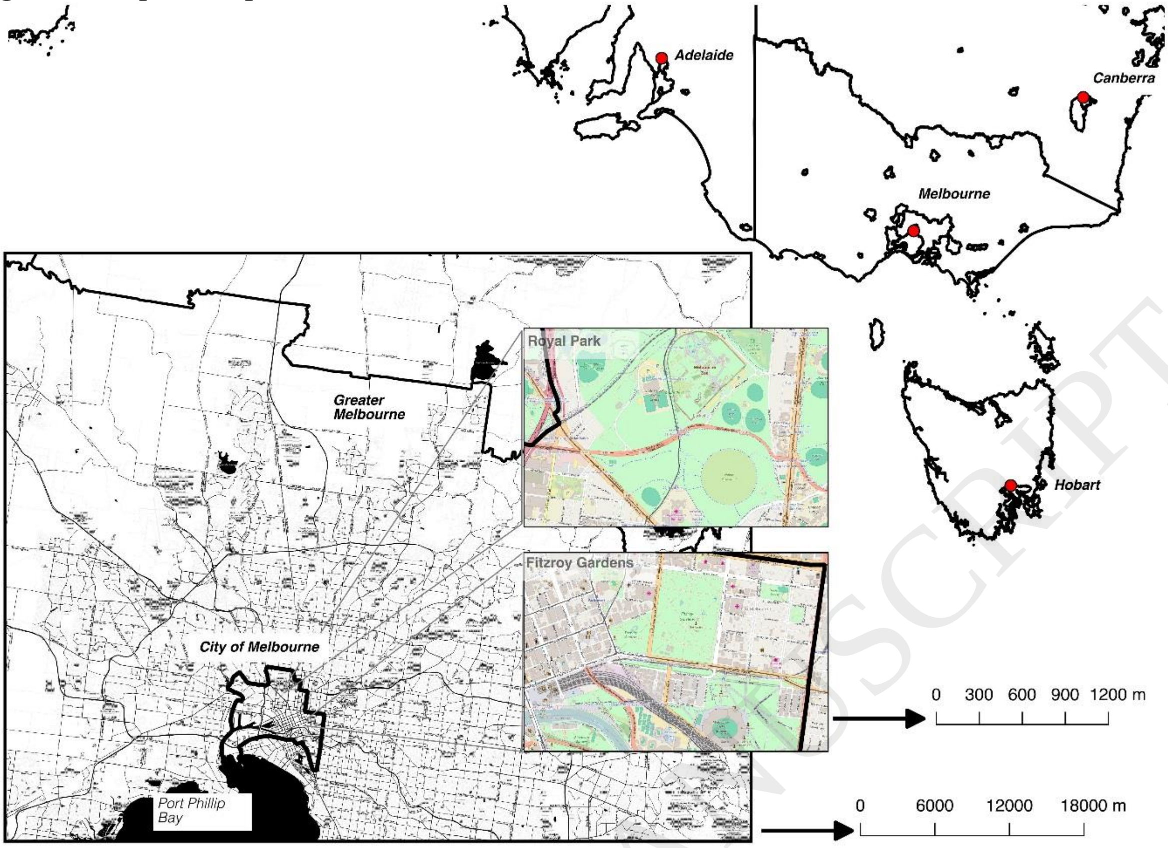


Figure 2: Still from Fitzroy Gardens video (a) with corresponding training image for the pixel labelling (b)

(a) (b)

Figure 3: Results showing one frame with the pixels labeled for Royal Park (a) and Fitzroy Gardens (b)

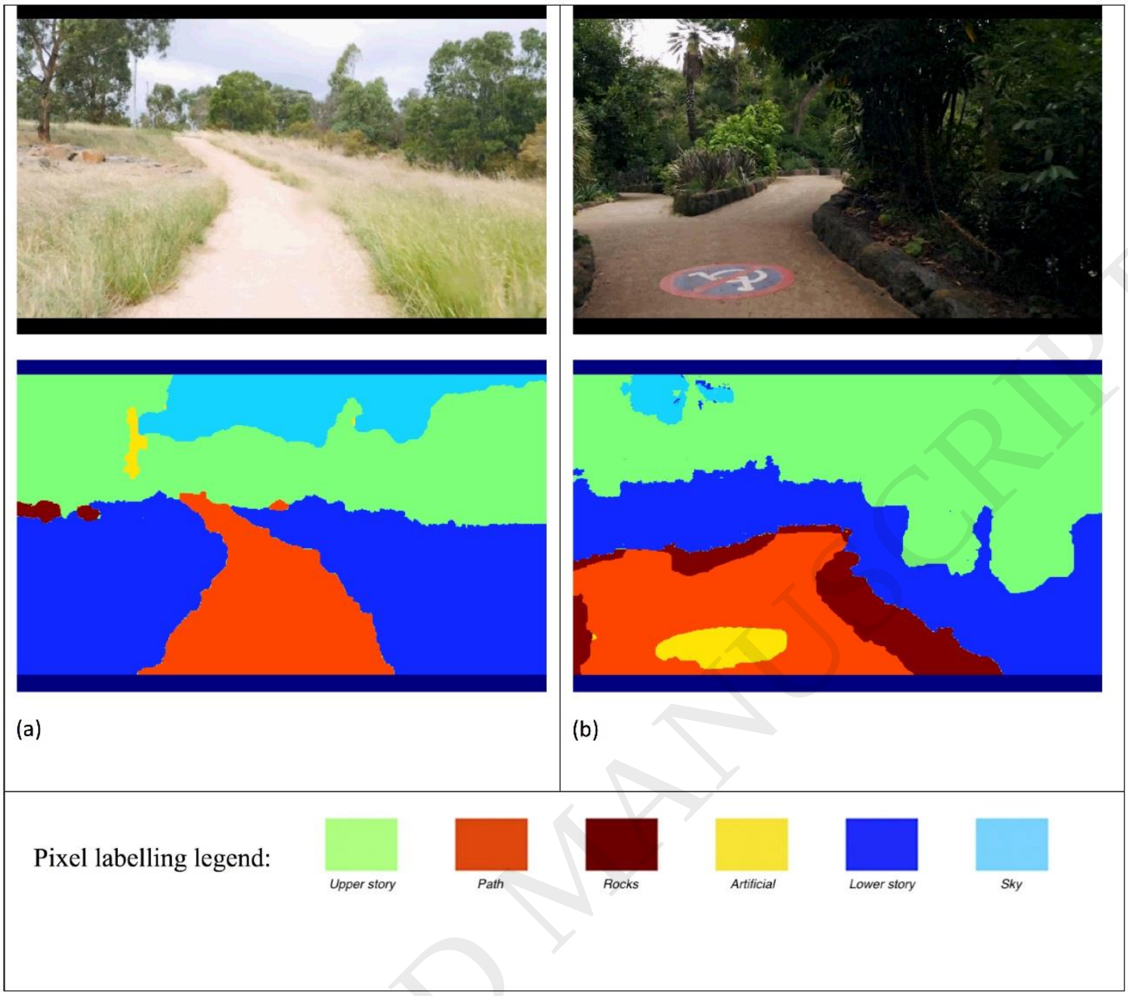


Figure 4: AOI in Royal Park (a) and Fitzroy Gardens (b) and dwell time for participant A on Royal Park (c) and Fitzroy Gardens (d) videos

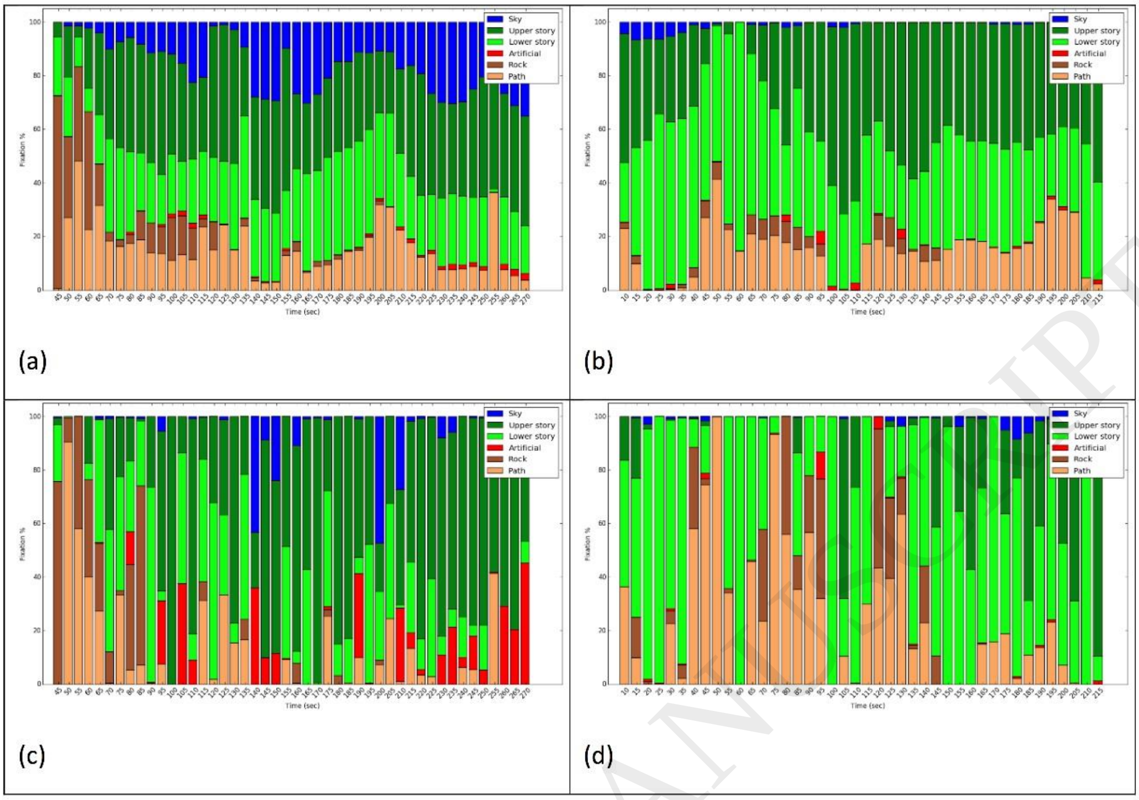


Figure 5: Boxplot of dwell time for viewed AOI categories for the two parks for the duration of the videos (3' 50'' seconds)

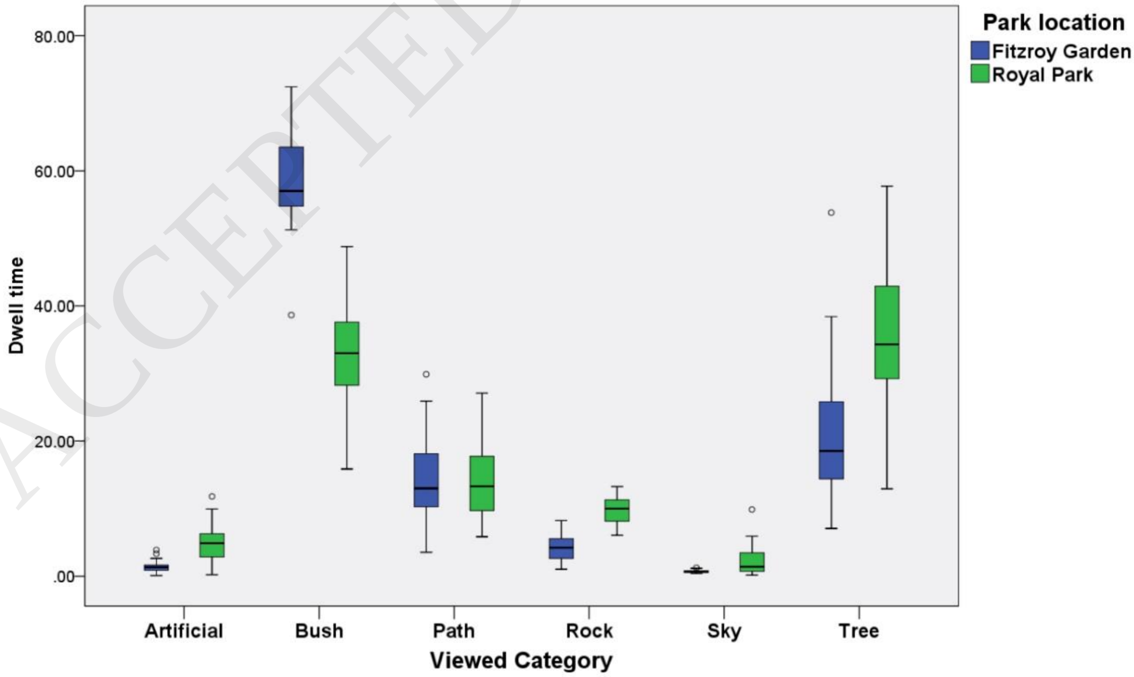


Figure 6: Boxplot of dwell time divided by the % surface areas of AOIs for each park for the duration of the video.

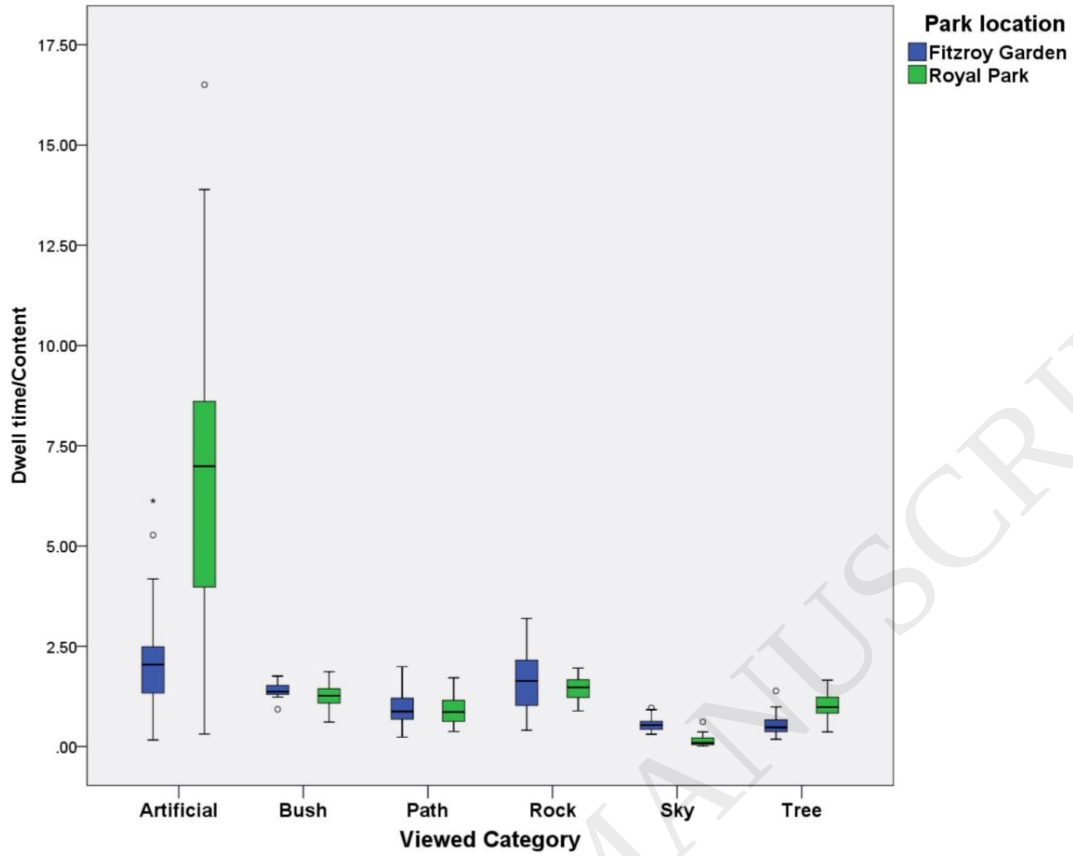


Figure 7: Examples of frames that relate to climbing stairs for Royal Park (a) and Fitzroy Gardens (b)

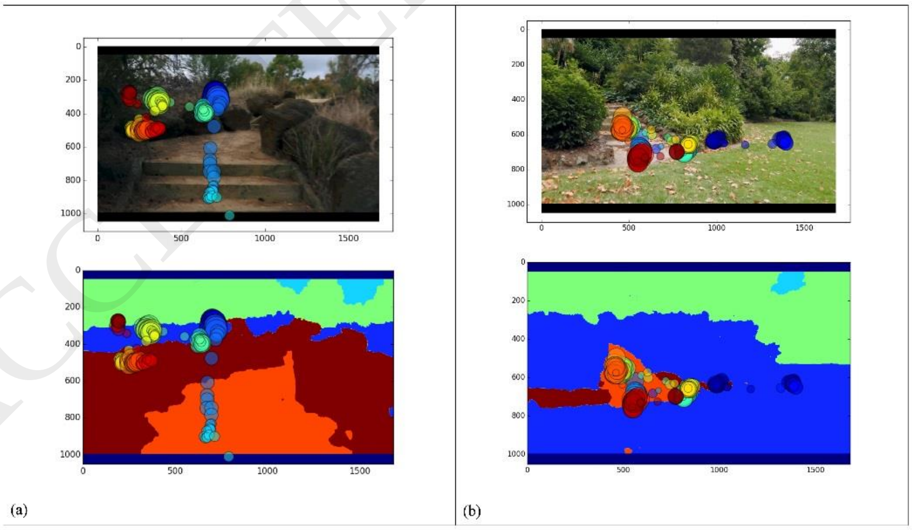


Figure 8: Attenuation of attention on the path during the step scene

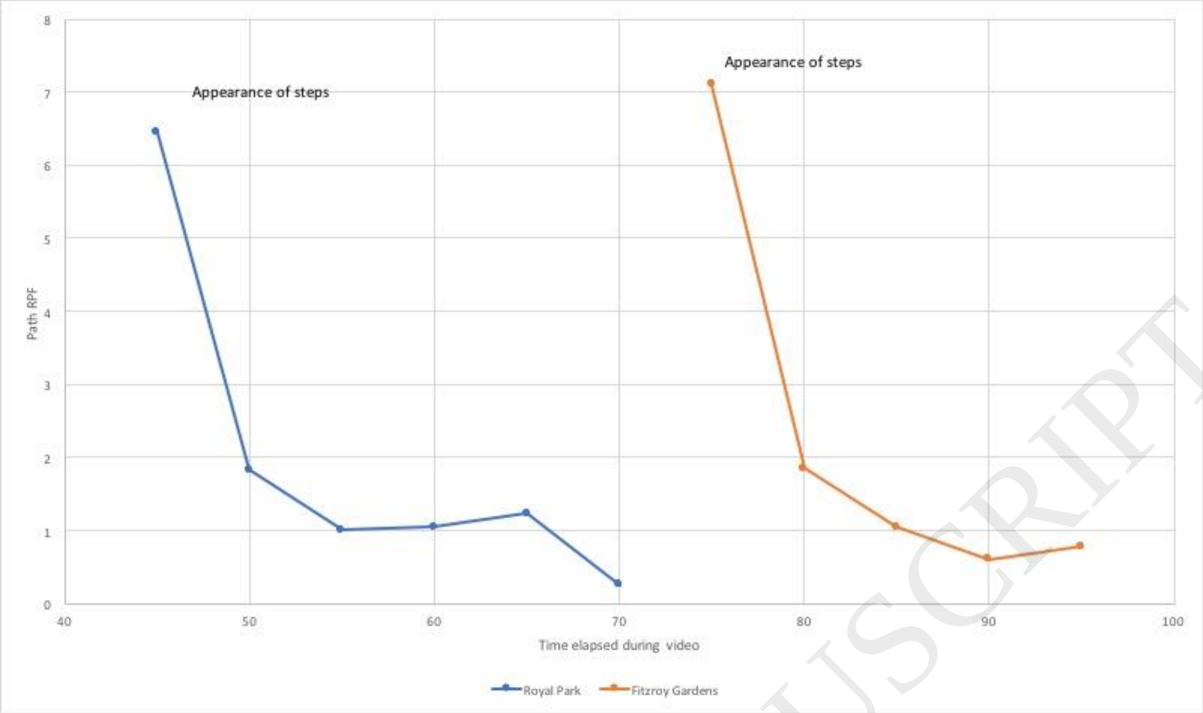


Table 1: demographic profile of the participants

Participant Gender	*N (%)*	*Ages*	*N (%)*	*Occupations*	*N*
Female	16 (46)	18-25	6 (17)	Students	16
Male	19 (54)	26-35	16 (46)	Retail	3
		36-45	10 (29)	University	2
		46-55	2 (6)	Hospitality	2
		N/A	1 (3)	Design	2
				IT	2
				Web design	1
				Mining	1
				Finance	1
				Engineering	1
				Government	1
				Legal tech.	1
				Printing	1
				Church	1
Total	35		35		35

Table 2: Description of areas of interest and features in the different videos

Area of interest	Royal Park example	Fitzroy Gardens example	Algorithmic properties of the machine learning that would enable the accurate identification of these
Upper-story vegetation	Trees, mostly Eucalypts e.g. *Corymbia maculata* Distant views and mostly of the whole tree	Trees, mostly exotic e.g. *Platanus x acerifolia*, *Ulmus procera*. Close up views. Sometimes only of the truck with some of the canopy	All vegetation in the top third of the scenery High internal contrast variation and edge orientation distribution. Colour.
Path	Fine gravel. Steps	Asphalt. Steps	Tan or black in colour. Occupies middle of the scenery. Homogenous internal texture and high contrasted boundaries.
Rocks	Granite boulders. Used to line the path. Occasionally features in the long grass.	Basalt cobbles used to line the path	Grey to black in colour. Lower half of image frame.
Artificial	Light poles, distant buildings, signage	Light poles, benches, signage	Shape. Contrast with surroundings
Lower-story vegetation	Path-side shrubs and lower vegetation. Tussock native grasses.	Path-side bushes and vegetation. Densely foliated garden beds. Short lawn grass and turf	Vegetation in the lower third. Grass. High contrast variability and edge orientation distribution.
Sky	Overcast sky	Overcast sky	Upper third of image. Homogenous local colour distribution with minimal internal texture.

Table 3: Comparison for Dwell time of view categories between two parks (Mann-Whitney U-Test) (FG: Fitzroy garden; RP: Royal park)

View Category	Park	N	M-W test Mean Rank	M-W test U	M-W test p
All Categories	FG	210	194.59	25,391.00	.007
	RP	210	226.41		
Artificial	FG	35	23.63	1028.00	.000
	RP	35	47.37		
Bush	FG	35	52.80	7.00	.000
	RP	35	18.20		
Path	FG	35	34.57	645.00	.703
	RP	35	36.43		
Rock	FG	35	19.31	1179.00	.000
	RP	35	51.69		
Sky	FG	35	25.77	953.00	.000
	RP	35	45.23		
Tree	FG	35	22.57	1065.00	.000
	RP	35	48.43		

Table 4: Comparison for Dwell time/content of view categories between two parks (Mann-Whitney U-Test) (FG: Fitzroy garden; RP: Royal park)

View Category	Park	N	M-W test Mean Rank	M-W test U	M-W test *p*
All Categories	FG	210	205.45	23,111.00	.394
	RP	210	215.55		
Artificial	FG	35	24.06	1013.00	.000
	RP	35	46.94		
Bush	FG	35	41.37	407.00	.016
	RP	35	29.63		
Path	FG	35	35.54	611.00	.986
	RP	35	35.46		
Rock	FG	35	37.57	540.00	.394
	RP	35	33.43		
Sky	FG	35	51.97	36.00	.000
	RP	35	19.03		
Tree	FG	35	21.37	1107.00	.000
	RP	35	49.63		

Table 5: Participant reactions to the parks

Royal Park rating	*N*	*Royal Park sample comments corresponding with the rating*	*Fitzroy Gardens rating*	*N*	*Fitzroy Gardens sample comments corresponding with the rating*
0-1.9	1	"Unkempt, spooky, too rural rather than metropolitan"	0-1.9	0	N/A
2-3.9	4	"Not much view"	2-3.9	1	"When we started the walk I felt like everything was really open and sparse. Very empty. As we continued the walk went a bit dark, really oppressive"
4-5.9	7	"Colours. I didn't like how it was brown and dried out looking"	4-5.9	2	"Variety and type of plants"
6-7.9	12	"Liked the footpath but didn't like the high dry grass – looked unkempt"	6-7.9	6	"Very covered from buildings and things like that"
8-9.9	8	"Liked that it was secluded"	8-9.9	19	"Rustic stairs at the beginning"
10	3	"I liked the native sort of bush land and the plants"	10	7	"Different trees. I love looking at different trees especially furry ones. And the different greenery"

Table 6: Spearman's rho Correlation Coefficients and Significances between visitors' park ratings and dwell time of view categories

	Fitzroy garden rating	Royal park rating
Tree	.486** (.003)	.299 (.081)
Path	-.316 (.064)	-.407* (.015)
Rock	-.271 (.115)	-.183 (.292)
Artificial	.162 (.352)	.274 (.111)
Bush	-.292 (.088)	-.187 (.281)
Sky	.245 (.156)	.337* (.048)

*. $P < 0.05$ (2-tailed). **. $P < 0.01$ (2-tailed). P values are presented in the brackets.