A Multimodal Measurement Approach Using Eye Tracking and Narratives to Investigate Visual Behaviour in Perceiving Naturalistic and Urban Environments

KhizarZ.Choudhry[a,b], Richard Coles[b], Robert Ashford[a], SalimKhan[a], Rabia R. Mir[a]

[a]Faculty of Health, Birmingham City University, Birmingham, UK.

[b]School of Architecture, BIAD, Birmingham City University, Birmingham, B42 2SU. UK.

Abstract: The majority of existing landscape research has been derived by conducting heuristic evaluations, without having empirical insight of real participants' visual responses. In this research, a modern multimodal measurement approach (using Eye tracking and Narratives) was applied to investigate visual behaviour in perceiving naturalistic and urban environments. This research is unique in exploring gaze behaviour on environmental images possessing different levels of saliency. Eye behaviour is predominantly attracted by salient locations. The concept of methodology of this research on naturalistic and urban environments is drawn from the approaches used in market research. Borrowing methodologies from market research that examine visual responses and qualities provided a critical and hitherto unexplored approach. This research has been conducted by using a mixed methods approach.

The results of this research corroborated existing landscape research findings, but they also identified potential enhancements to the published research itself and the methodology adopted in the previous research. This study focused on initial impressions of environmental images with the help of eye tracking. Taking into consideration the importance of the image, this study explored the factors that influence initial fixations in relation to expectations and preferences. This research contributes both methodologically and empirically to Human-Environment Interaction (HEI). In terms of key findings, this research identified that each participant has a unique navigation style whilst surfing through different elements of landscape images. This individual navigation style is termed the *'Visual Signature'* which in this study is defined as an individual pattern of fixations which emerge as a result of involuntary initial response within the first 3-5 seconds after exposure to a naturalistic or urban image.

Keywords. Human-Environment Interaction (HEI), multimodal measurement, eye tracking and narratives

Background:

To date, little empirical work has been carried out in the area of eye tracking and environmental research. This research has taken an empirical, human centred approach to develop further existing research in the environment with the use of new research techniques such as eye tracking. The effects of environmental change and development on people and on their attachment to natural areas have gone relatively unstudied [1, 17]. Therefore, within this study, the 'impact attributes' of surrogate landscapes on the individual through exposure to images derived from naturalistic and urban environments was undertaken utilising eye tracking and interviews . This study set out to bridge the gap in existing environmental research by bringing in new data collection (eye tracking) and analysis techniques in the field of environmental research. This was with a view to potentially revealing the characteristics of green infrastructure components that may have an impact on physical, social and cognitive health functioning and well-being. Thus, it is the papers premise that meanings and behavioural values need to be examined through behavioural responses of urban residents using more complex methodologies to elicit individual's preferences for an environment. This study's finding has generated new evidence on the qualities of green infrastructure in towns which has the potential to help improve the planning and design of green infrastructure in towns. It is also suggested these results need further exploring and a larger study is required to validate this pilot work.

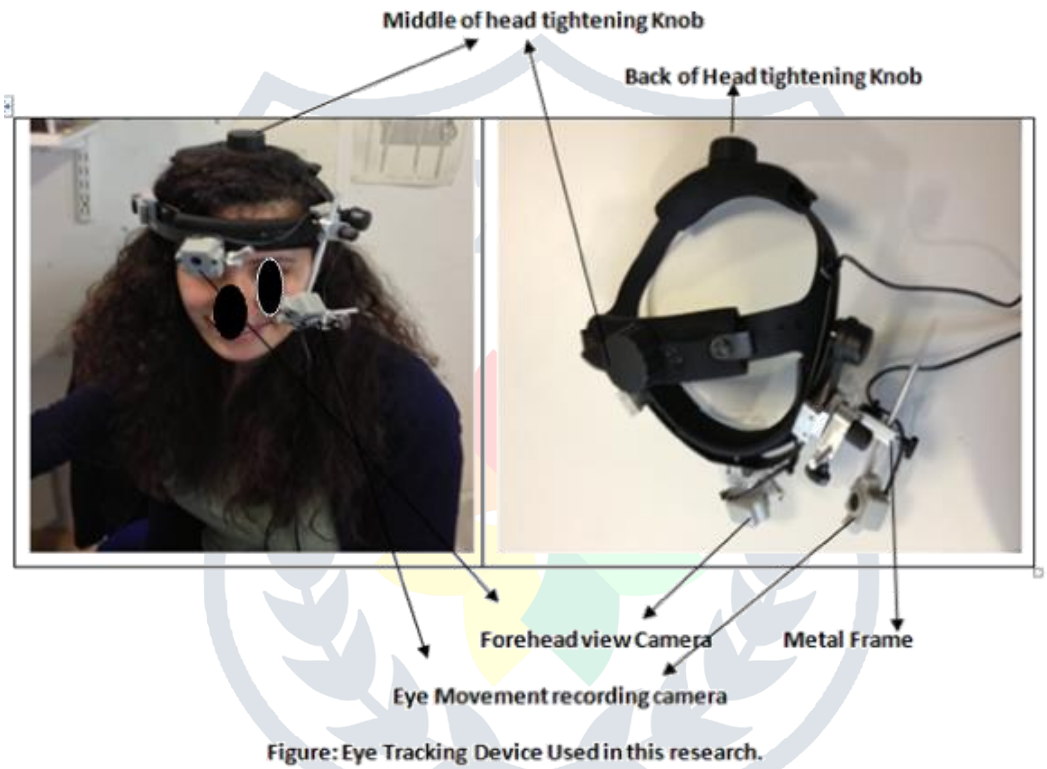
It has been suggested that it is important to gain a knowledge of the way people observe and perceive landscapes, and this should be incorporated into the way in which landscape planning and management are conceptualised. [1, 17]. So far, different landscape perception paradigms have been formulated [2] and analysed using questionnaires and in-depth interviews. The most frequently used stimuli in such empirical research are photographs or in situ observations [6, 5, 4, 3, and 10]. As alluded to earlier, an objective way to measure individual's observations of landscapes, is to use eye tracking techniques. These techniques can be used to record the speed and direction of eye movements (saccades) and the position and duration of fixations whilst observing images. Eye tracking measurements have been extensively studies in the field of psychology [7, 10]. Other disciplines have also utilised this technique and there is some excellent examples in geography [7, 10] and landscape science [9, 8]. Landscape photographs are often used in landscape perception research [10]. It is believed that, coupled with eye tracking, a more in-depth exploration of landscape perception may be possible. According to Dupont (2014) a better landscape management and planning can be achieved by bringing in people's perspective, up to recently the knowledge and insight, about how people perceive landscapes is gained through questionnaires and in-depth interviews [17, 18].

Methods:

The aim of the study was to determine the properties and attributes of green infrastructure and its land use components through preference by eye tracking experiences whilst participants were exposed to a range of environmental images. This allowed for an investigation into the impacts of green infrastructure (independent variable) towards people's physical, cognitive and social health and well-being (dependent variable).

A video-based eye-tracking system was used to project harmless infrared light onto the participants eyeball and track reflections from landmarks of the eye. These were recorded by two cameras fitted with Image processing software (Gaze Tracker™) and were subsequently used to identify and map eye positions to the display (Figure1). Infrared light was absorbed by the eye at < 1% of the Maximum Permissible Exposure Level as certified by the American Standards Institute (ANSI Z 136.1-1973).

There are numerous video-based systems available for use. This study utilised a Sensori Motoric Instruments (SMI Sensori Motric, 1999) head-mounted system (HED-II) which provided a sampling rate of 10 Hz, thus correcting for head and general movement (Figure 1).



Figures 1. A Sensori Motoric Instrument head mounted system

Current available headgear systems were deemed inappropriate for the study due to set up times, additional hardware being required and several systems being overly obtrusive for the participant. Therefore, a purpose built eye tracking device was used in this research which required minimal time for set up; was less intrusive; and allowed a full range of head movements during exposure whilst still measuring pupil size. The organisation and analysis of the data was facilitated by a Gaze Tracker™ software package.

The SMI's Head Mounted Eye Tracking Device (HED-II) uses two small cameras (eye camera and scene camera) with an infrared light mounted on an ophthalmoscope band resulting in a combined total weight of 450g (Figure 2). A piece of glass is strategically placed which reflects the downward pointing light into the eye. The front of the eye is illuminated with the infrared light. This produces a dark pupil and a corneal reflection. The scene camera (forehead view camera) records what the participant is viewing on the screen. This recording and the eye movement data are transferred via video encoding hardware to a laptop enabling the creation of mpeg video files of the scene video with an over-laid moving dot representing the participant's eye movements. The laptop is enabled with iView™ software and is connected to the eye tracker to analyse the transmitted data from environmental images; the Point of Regard is then computed (Sensorimotor Instruments, 1999). The generated eye movement data is subsequently exported to Excel spreadsheets for further analysis.

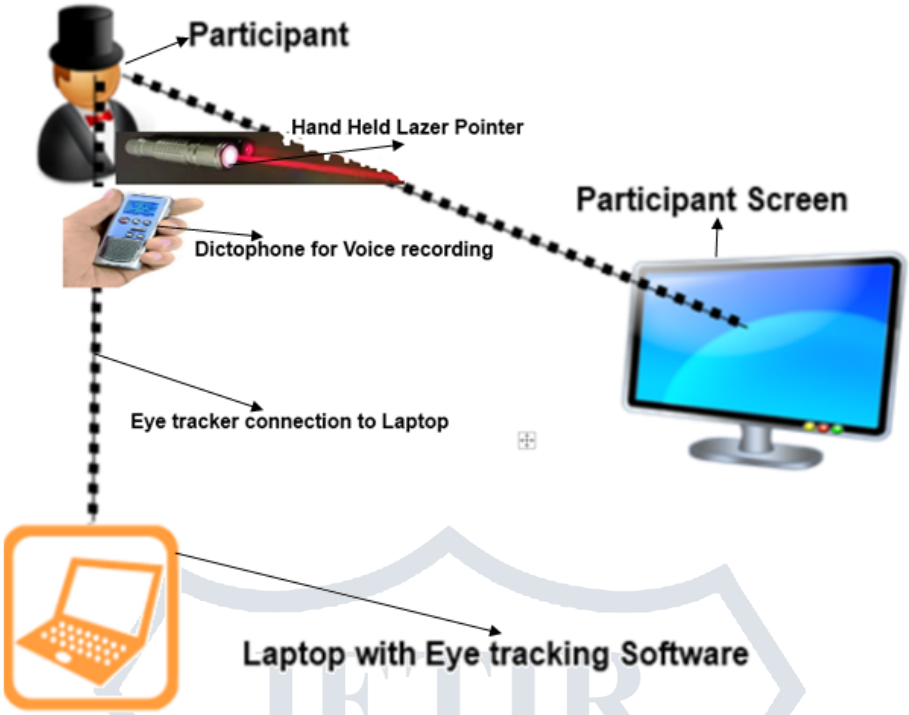


Figure 2. Overview of lab setup for data collection of study

Eye movements were recorded using an eye tracking camera capturing eye movement data at a sample rate of 10Hz. An adjustable camera was placed in front of the participant's eye without obscuring the participants' view. The presentation of the stimuli was controlled by means of the Gaze Tracker™ software and presented on the screen viewed by participants from a distance of 6 feet from a large flat screen (60'x 60' inches). The GazeTracker™ software also recorded the eye-movements and enabled the researcher to view the data. ASL Eye pose software also recorded the data and was used as a back-up. A detailed presentation of the set-up is presented in Figure 3.

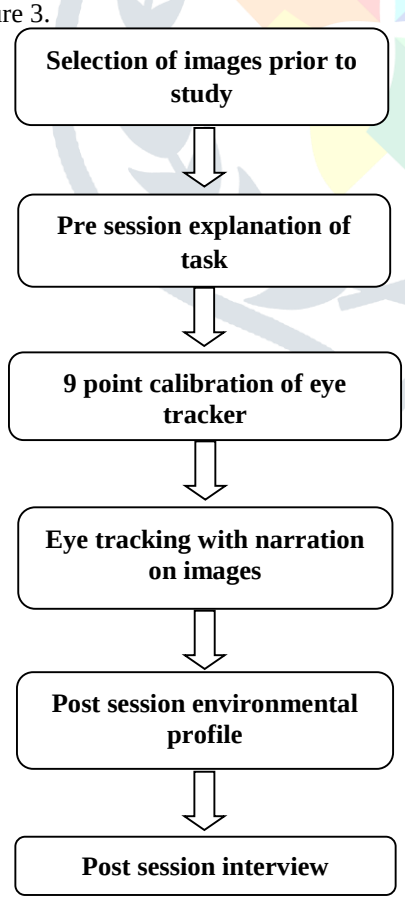


Figure 3. Collection of evaluation techniques as applied across this research

As this was an exploratory study, the need to control viewing conditions was of pivotal importance. Hence, all experimental work was undertaken in a temperature controlled room with no windows. Screen size (60x60 inches), seating distance from screen and height of seat were all kept constant.

Images were selected from a scene classification project data base run by Stanford Vision Laboratory, UIUC Beckman Institute. Full permission was granted for use of such images. Selection of images of natural and urban landscapes and subsequent categorization of good and bad examples of natural scenes was done via the Delphi technique* (Mullen 2000). This technique uses a consensus approach to identify which images are highly positive and which have negative associations. This approach allows for a description of an image and the rationale for its selection. A maximum of 12 images subdivided into three broad categories; pure natural environments, heavily urbanised environments, and environments containing both urban nature/natural elements were taken forward for use with participants and eye tracking. Selection of the images from the wider Stanford Vision pool of images was facilitated by asking volunteers (as naive subjects i.e. non landscape professionals)to rate images as positive, negative or neutral categories using Hartigs's (1991) perceived restorativeness scale. Images were rated on a 5-point scale by asking participants to complete the phrase 'it makes me feel.....'(Figure 4).

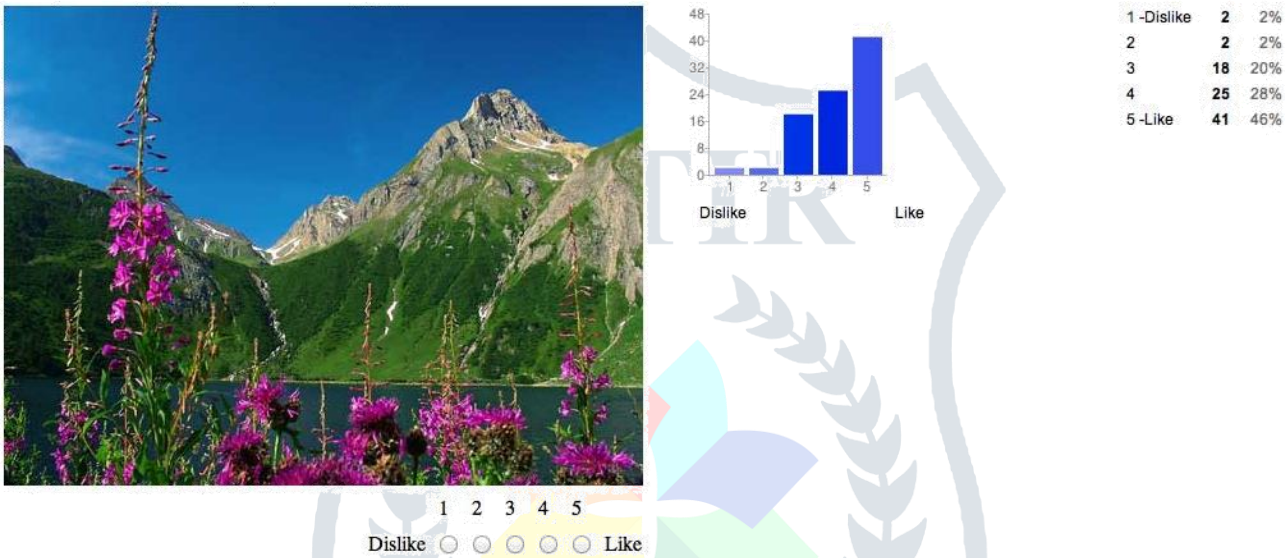


Figure 4: Example of five point Likert scale and its results.

Once participants had been briefed on the process and equipment, a 9 point calibration[1]was conducted whilst the participant's eye movements were recorded. Upon completion of this, the eye tracking session commenced (Figure 5).

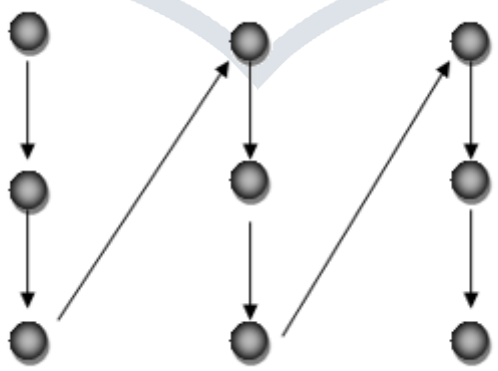


Figure 5. Gaze movements during nine point calibration on screen.

[1] Nine point calibrations: Calibration of eye tracking equipment was conducted for each participant by asking each individual to look at 9 different points on the screen one at a time.

Each participant looked at each of the twelve images individually and provided a running narrative that was recorded, and later transcribed. Upon completion of eye tracking, a short semi structured interview was conducted with each participant to test recall and recognition of exposure to the images. In addition, a Connectedness to Nature Scale (CNS[2]) questionnaire was also undertaken to elicit user profile (e.g. name, age, gender); previous environmental experience (e.g. frequency and purpose of environmental use); and a measure of an individual's trait levels of feeling emotionally connected to the natural world.

This study takes the view that quantitative and qualitative approaches complement each other. Although there has been much debate about the dichotomy between quantitative and qualitative paradigms [7]. Research can be a pragmatic combination of quantitative and qualitative methods [11], in which the findings from each technique shed light on the findings from the other. Such 'triangulation' is used deliberately within studies involving participants to explore the relationship between experience and behaviour [8]. De Lucio (1996) suggest that the combination of eye tracking and other evaluation techniques gives a richer set of data. The primary evaluation technique used in this research as an indicator of visual behaviour is eye tracking. Eye tracking data is complemented by data from narratives, background questionnaires and post-session interview.

Results:

This study identified that the initial 3-5 seconds of viewing an image is of extreme importance in describing the scanning behaviour of individuals and is referred to in this paper as the '***Visual signature'***. Visual signature[1] formation is an important finding from this study. Previous research identified that eye gaze patterns can be used to classify the type of task the user is performing [11]. However, findings from the present study identified that during eye tracking data analysis, individuals show two types of responses when exposed to a naturalistic image, an immediate involuntary response (within the first 3-5 sec) and a delayed response (according to the image being liked or disliked). Hence, there is a need to examine the way participants look for specific design elements such as which sections of images are more prone to first fixations. By capturing the participants' initial eye movements when they visit an image, it is possible to identify which location on the image attracts immediate attention and how the image tends to be scanned (*visual signature*). This can be related to what information is presented at those locations on the image and whether immediate attention is drawn to important information (points of interest) (Figure 6, 7, 8).

Figure 6 Immediate Visual Response: Emergence of individual scanning behaviour immediately after initial exposure.

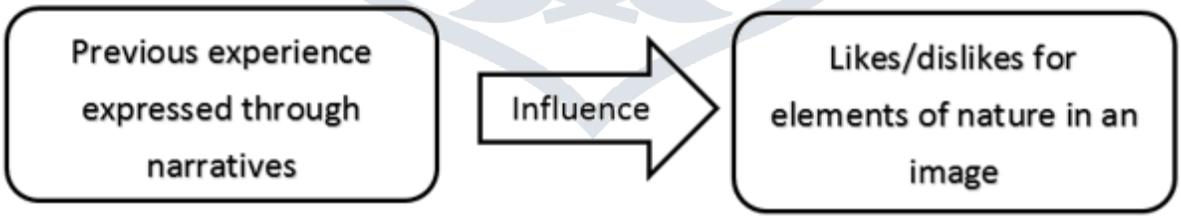


Figure 7. Delayed Visual Response: Influence of previous experience and exposure expressed through narratives.

When exploring the influence of previous experiences on visual performance on landscape images, the observed results from this study suggested:

Expectation and first impressions: In most cases, participants looked at the centre of the viewed image for quick clues as to the kind of environment being shown. This may be influenced by the participant's previous experiences of interacting with such an environment. The eye tracking data and accompanying narratives made clear that participant's decided within the initial few seconds of exposure whether the observed image was liked or not.

[2] The back ground questionnaire is a modification of a questionnaire used by Mayer etal., 2003.

Participant adaptability: Participants 'learn' where to look over repeated exposures. Interestingly, the effect of their previous experience over rules their quick adaptation to a specific type of environment. Some participants who were exposed to images prior to the eye tracking exercise focused on the centre of the image and did not move to different quadrants of the image.

Importance of task: The purpose, relevance and familiarity of the task itself influence users' behaviour. For example, some participants were more focussed on the study from inception to completion. This was more apparent in participants who obtained higher nature profile scores. Such participants spent more time on each image and tended to scan each image with greater detail.

Pattern of Gaze: Eye gaze patterns can be used to classify the type of task the participant is performing. The results show that each participant has a unique visual signature of eye movement (Figure 5). In addition to the participant's unique signature, the eye tracking data provides an insight into whether a participants search behaviour is 'exploratory' or more focused. "Previous research suggests that a system can be developed for classifying user tasks based on prior knowledge of the categories of certain tasks a user typically performs and training sets of eye movement data for those tasks [11]." In this study, data from participants who were exposed to images prior to the eye tracking exercise showed more 'exploratory' patterns, whilst, participants who were new to the task showed a more 'well focused' pattern.

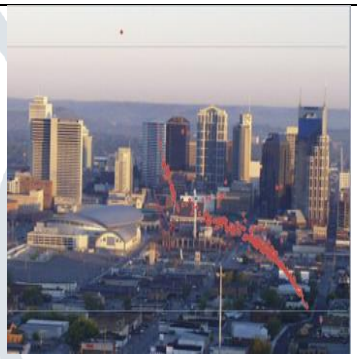
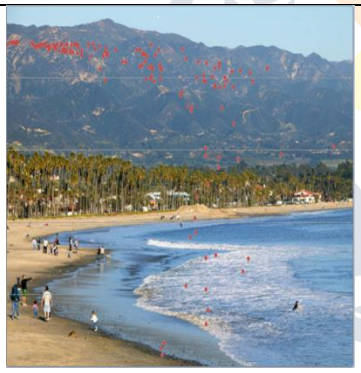
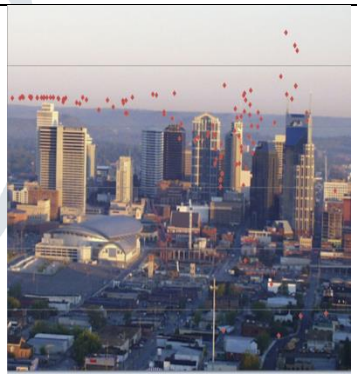
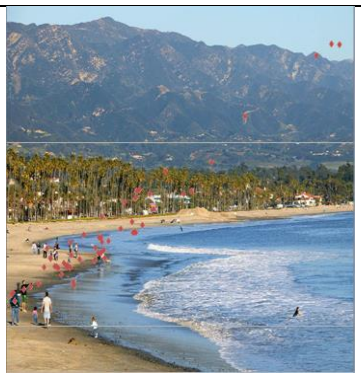
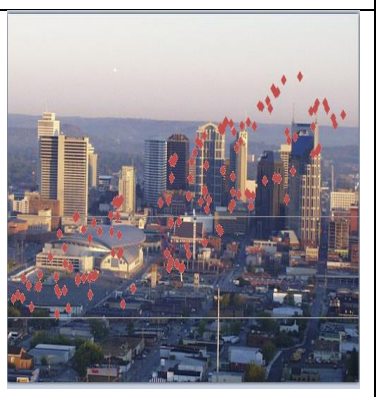
	Image 1	Image 2	Image 3
Participant A			
Participant B			
Participant C			

Figure 8. Pattern of visual scanning and the Visual Signature

Conclusion:

The application of a mixed methods approach revealed aspects of landscape research that had not been addressed by existing methods of landscape research. The detailed information of shifts of visual search behaviour on the stimuli landscape images provide insight of which design elements attract attention and where the participant's astronomically look when viewing specific landscape images. The majority of existing landscape preference research has been derived by personal reviews and anecdotes without reporting the involvement of human participants. This study tested and elaborated the existing landscape research by capturing visual search behaviour by use of eye tracking and narratives. The findings are expected to benefit human environment researchers as they will learn about the potential use of eye tracking for evaluations of landscapes.

The stimuli used in the study were static environmental images over a very short period of time. The participants were required to look at single environmental images in a sequential manner, each from a different type of environment, and were therefore not allowed further interaction within any other type of environment at that time. Participants reacted with the environment image, looked at different things and when satisfied moved to the next image. Future research should explore the effect of more complex tasks across multiple environment images of different environment set ups shown at the same time. In addition, a more friendly and realistic environment (virtual realistic smart room) might be used in the future but the complexities of measuring eye movements such an environment are fully acknowledged here. This study will also refine existing eye tracking methodologies used in visual behaviour research for example Autism and Asperger's research.

REFERENCES:

[1]. Ryan, J. D., & Cohen, N. J. The nature of change detection and on-line representations of scenes. Journal of Experimental Psychology: Human Perception and Performance, 30(5), 988-1015. (2004)

[2]. Scott, Karen E & Benson, John F (2002). Public and Professional Attitudes to Landscape: Scoping Study. Scottish Natural Heritage Commissioned Report F01AA303, pp 51, SNH, Edinburgh. (2002)

[3]. Andrew L, Walker and Rosalind, Kimball Moultan, photo album: Images of Time and Reflection of self ,Kimball,Qualitative Sociology;Summer89, Vol. 12 Issue 2, p155 (1989)

[4]. Palmer, J.F., 2004. Using spatial metrics to predict scenic perception in a changing landscape: Dennis, Massachusetts. Landscape and Urban Planning, 69 (2/3), 201-218. (2004)

[5]. Ode, A., Tveit, M., & Fry, G. Capturing Landscape Visual Character Using Indicators: Touching Base with Landscape Aesthetic Theory. Landscape Research, 33(1), 89-118. (2008)

[6]. Hagerhall, C. M. Clustering predictors of landscape preference in the traditional Swedish cultural landscape: Prospect-refuge, mystery, age and management. Journal of Environmental Psychology, 20, 83–90 (doi:10.1006/jevp.1999.0150). (2000)

[7]. Hammersley, M. 1998. 'Get real! A defence of realism', in P. Hodkinson (ed) The Nature of Educational Research: Realism, relativism, and postmodernism. Manchester: Manchester Metropolitan University.

[8]. Mackay, W. E., & Fayard, A.-L. (1997). HCI, natural science and design: A framework for triangulation across disciplines. Proceedings of DIS 97, 223-234. New York, NY: ACM.

[9]. Mayer, F. S., Frantz, C. M., Norton, C., & Rock, M. (2003). Self-awareness and connectedness to nature. Paper presented at the American Psychological Society, Atlanta.

[10]. Nevalainen S., Sajaniemi J. (2008) Experiment on the Short-Term Effects of Engagement and Representation in Program Animation. Journal of Educational Computing Research 39(4), 395-430

[11]. Ooms, K., De Maeyer, P., Fack, V., 2010b. Analysing Eye Movement Patterns to improve Map Design. Proceedings of the 18th International Research Symposium on Computer-based Cartography and GI Science (AutoCarto 2010): Geospatial Data and Geo visualisation IAPS (2012)

[12]. Savenye, W.C. and Robinson, R. S. (2012). Using qualitative research methods in the learning sciences. In Seels, N.M. (ed). Encyclopaedia of the Sciences of Learning. Heidelberg Germany: Springer.

[13]. Tveit, M.S., Ode, A., Fry, G.,Key concepts in a framework for analysing visual landscape character. Landscape Res. 31, 229–255. 2006.

[14]. De Lucio, J. V., Mohamadian, M., Ruiz, J. P., Banayas, J., Bernaldez, F. G.,. Visual Landscape Exploration as Revealed by Eye Movement Tracking. Landscape and Urban Planning, 34, 135-142. (1996)

[15]. Sevenant M., Antrop M.. Trans disciplinary landscape planning: Does the public have aspirations? Experiences from a case study in Ghent (Flanders, Belgium), Land Use Pol.,27(2), 373–386. (2010)

[16]. Iqbal S.T., Zheng X.S., and Bailey B.P. Task-Evoked Pupillary Response to Mental Workload in Human-Computer Interaction. Proc. CHI'04. Vienna: ACM, 1477-1480. (2004)

[17]. Dupont L, Antrop M & Veerle Van Eetvelde (2014) Eye-tracking Analysis in Landscape Perception Research: Influence of Photograph Properties and Landscape Characteristics pages 417-432 Geography Department, Ghent University, Belgium Routledge (2014)

[18]. Scott, Karen E & Benson, John F (2002). Public and Professional Attitudes to Landscape: Scoping Study. Scottish Natural Heritage Commissioned Report F01AA303, pp 51, SNH, Edinburgh. (2002)