Measuring Perception of Urban Design Elements in Virtual Environments Using Eye Tracking: Benefits and Challenges

Phillip Fernberg[1], Emily Tighe[2], Morgan Saxon[2], Charisse Spencer[1], Scott Johnson[1], Jeanine Stefanucci[2], Sarah Creem-Regehr[2], Brent Chamberlain[1]

[1]Utah State University, Utah/USA · phillip.fernberg@usu.edu
[2]University of Utah, Utah/USA

Abstract: Virtual reality offers rapid production of alternative design scenarios and exploration of design options, providing the designer extensive control over elements and experience. Eye tracking in VR provides a means to measure user perception of visual elements objectively and thus, could be useful in the landscape and urban design process. However, there are few studies that showcase an application of eye tracking in 3D for this purpose. This paper provides a case study that implements eye tracking in VR-based design assessment and discusses benefits and challenges of implementation. Benefits include a more structured, discretized design process and greater empirical controls for design assessment. Challenges include the technological implementation and the evaluation process to draw meaningful associations between the data and design decisions. This work will contribute toward future empirical research to explore design process, measure design scenarios and understand the usefulness of novel technologies for evaluating these designs.

Keywords: Virtual reality, eye tracking, gaze tracking, urban design

1 Introduction

Landscape architecture and urban design knowledge production employ a variety of theoretical approaches, both empirical and artistic (CARMONA 2021, DEMING & SWAFFIELD 2011). Each approach offers differing benefits, with the empirical approach providing a means to structurally and quantitatively deduce how various factors affect performance of design spaces and their elements. Such quantitative evaluations are informed by data analysis techniques, sensors, and technology, as well as through immersive virtual reality (VR) (EMO et al. 2016, EWING & CLEMENTE 2013, MORELLO & RATTI 2009, WALLER 2005, F. ZHANG et al. 2018). Furthermore, approaches such as 3D isovists (MORELLO & RATTI 2009), computational revisiting of foundational concepts like imageability (FILOMENA et al. 2019), and blending empirical social observations with spatial analytics facilitate a more quantifiable understanding of how some urban design decisions shape and mediate the experience of the spaces they produce (RAJENDRAN & ODELEYE 2020). In a search for more objective measures that could replace or augment aspects of subjective measures (e. g. surveys) of spatial design, we apply established methods of eye tracking from other disciplines (cognitive sciences in particular) to experimental research in order to understand how eye tracking in 3D VR environments may benefit the landscape and urban design process. For this paper, we reflect on the following questions: 1) How to implement eye tracking in VR-based design assessment, and 2) what are the benefits and challenges of this implementation? We offer this reflection to facilitate discussion about the role of the technology within landscape architecture and urban design.

Journal of Digital Landscape Architecture, 7-2022, pp. 463-470. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-724-3, ISSN 2367-4253, e-ISSN 2511-624X, doi:10.14627/537724045.

2 Implementation of Eye Tracking in VR

2.1 Why VR? Why Eye Tracking?

The importance of VR as a research medium in this context is critical. An immersive virtual world developed using gaming engines mediates the need for both control and flexibility in the environment. It can evoke a sense of place that affords visitors feelings of agency and spontaneity similar to the "real world", while also retaining controlled ways of measuring those visitors' objective and subjective responses to the place itself, thus meeting the research needs of both urban designers and cognitive scientists. There is a large body of research demonstrating the effectiveness of VR for conveying real world experiences such as in educational settings (HAMILTON et al. 2021), rehabilitation (PEREZ-MARCOS 2018), and exposure therapy for phobias (CARL et al. 2019). Contemporary head-mounted displays support stereo viewing, wider fields of view, head tracking, and high-fidelity graphics, all contributing to "presence" and increased utility for many applications (SLATER 2018). This work provides a basis for the use of VR in urban design. Many studies have shown that eye tracking behaviours are quite similar when comparing performance in VR to the real world. For example, LI et al. (2016) found that visual search patterns were similar across naturalistic 2D and 3D immersive environments in that attention was increasingly allocated to relevant regions for identifying targets during visual search in both settings.

Eye tracking in VR has been used in a few architectural-related studies (XIANGMIN et al. 2021, L.-M. ZHANG et al. 2019), but most eye tracking landscape studies to date have used 2D static images (DUPONT et al. 2016) or animated simulation on a monitor (REN & KANG 2015). Eye tracking is used in many fields and can provide measures of fixation, order of fixation, and pupil dilation. In psychology studies these measures have been used to make claims about visual attention and arousal (KIM & LEE 2021), by elucidating which objects in the environment the viewer attends to, which strategies the viewer employs while searching for target objects amongst distractors, and whether some objects in the environment are allocated more attention than others. Eye-tracking allows for more objective measures of perception and attention compared to interviews or questionnaires (DUPONT et al. 2014). Taking it a step further, through eye tracking it has been shown that VR space can be simulated accurately to the point that there is an overall greater correlation of eye fixations between the real physical architectural space and the virtual reality space than there is between the real physical architectural space and the space communicated through plan and section drawings (HERMUND et al. 2018). This suggests that through both eye tracking and VR we could more objectively measure visual responses to designed spaces, such as salience, attention, memory, and navigation. Still, further research on the effectiveness of eye tracking as a tool for spatial design research is needed, particularly on whether it can help us investigate the interplay between design intent and perception in urban spaces (i. e., what the designer intends the user to perceive and what they actually perceive when interacting with the space).

2.2 Case Study: A Long Road and an Architectural Smörgåsbord

The inspiration for reflection on eye tracking in VR comes from current research endeavours to identify scientifically centred means for evaluating design. Our approach leans heavily on spatial cognition, wherein we focus on connecting what people say they experience (through subjective surveys and recall tasks) and what they objectively experience (e. g. through eye tracking). Our multi-disciplinary workflow draws from computer science, cinematic and gaming design, and environmental psychology coupled with traditional urban design research instruments. The purpose in these explorations is twofold. First, we investigate several urban design metrics and how to embed them in an empirically driven placemaking process. We do so by utilizing 3D gaming engines and licensed prefabricated 3D assets to create a set of procedurally generated urban environments which can be explored in VR. Second, we employ these immersive virtual environments (VEs) to test the effects of spatial design decisions on human perception. For the reflective purposes of the present paper, we will use an example from one of our pilot studies in which we constructed a virtual experience on a single road wherein subjects interacted with multiple vernacular environments at various scales and densities then subjectively registered any spatial changes they perceived (see Figure 1). This study was our first to use integrated VR eye tracking tools and is a good primer for rumination on the prospects and problems with adapting the technology for landscape and urban design-specific research purposes.

Using a head-mounted virtual display, participants navigated down a virtual road at a set speed and provided a response, via an Xbox controller, where they felt the environment changed. Participants also provided a verbal response, on a scale from 1 (small change) to 5 (large change), to indicate the magnitude of the change and answered debriefing questions to elucidate what elements of the environment, if any, they used to inform their responses. While the behavioural responses and debriefing responses alone provide information about whether and how individuals detect changes in their environment, the eye-tracking data will help us determine which design elements participants used to determine whether a change occurred and whether and how their attention allocation changed as they progressed through the experiment.

The implementation of eye tracking was developed using the VIVE Pro Eye, one of the few HMDs with integrated eye tracking hardware and a software development kit (SDK). The SDK plugin interfaces between the gaming engine (Unity) and the headset. Eye tracking uses an equivalent of a ray trace that can record all times when the ray intersects with a given object. Our implementation uses a ray which is a narrow conical projection (think narrow ice cream cone) extending from the eye to the furthest perceivable distance in the scene. The conical projection is intended to align with the *fovea centralis* of the eye rather than a laser-beam-like projection (see Figure 2, depicted as a green area). When the projection collides or exits a game object these data are recorded (often in milliseconds). Game objects can range from small elements, such as a window on a building, to the building itself, as well as any surface element(s). Our initial efforts used the building itself. The data we have collected show a vast array of collisions and exits, which leads to a significantly large dataset. Data collected include a timestamp, the object name, centre coordinates of the object, and an angle between the HMD and the gaze ray (e. g., 0° means the gazer was looking straight). The timestamps can be used to derive the time length of gaze fixation on a certain object. From the coordinates, the distance between the player and the object can be figured then recorded.

Fig. 1: VR environment used for pilot study

Fig. 2: Example rendering of eye tracking in VR to show object identification (building) through focal beam (depicted as a green area)

From these data it is possible to determine total dwell time (TDT) and total fixation count (TFC) for specific areas of interest (AOIs). AOIs are regions in the environment that are important to the experimental design (HOLMQVIST et al. 2011). TDT refers to the amount of time participants spent looking at a particular object or AOI. Total fixation count indicates the number of fixations on an object or AOI. TDT can be calculated using object collision and exit timestamps, and TFC can be calculated using the number of object collisions. Higher TDT and TFC indicate increased visual attention and suggest that specific elements or areas are more important and interesting (HOLMQVIST et al. 2011). Understanding the relationship between TDT and TFC associated with a given object allows us to explore the geospatial and temporal pattern of space observations and can provide clues about what individuals may infer about the environment.

2.3 A Note on Coding Conventions for Embedding Eye Tracking in 3D VEs

To produce TFC and TDT, we needed to identify an AOI (each building) and have a means to consolidate each eye tracking event. To accomplish this, we needed to develop a 3D VE, a module to produce eye tracking events and a script to consolidate the data into meaningful metrics. The environment was coded using block-scale prefabs, so a prefab could be loaded in real-time and repeated in any random order. This development necessitated the curated placement of objects in specific locations. The rapid development was supported by copying and rotating or moving buildings to different locations on the block. What we recognized later, is that the copying of buildings became a challenge for producing meaningful data because object names could be repeated. This made it difficult to determine metrics such as TDT and TFC because these metrics should be associated with a unique object. Thus, in future developments it will be imperative to rename any copied object uniquely. Additionally, a suffix should be added based on the incremental order that a prefab is loaded.

3 Discussion: Benefits and Challenges of Eye Tracking

The current case study, along with previous pilots and our ongoing work, bring to view a range of potential benefits and challenges for implementing eye tracking in VR-based design research. We will focus on two benefits and two challenges encountered in the initial studies. The first potential benefit worth noting is a more structured way of discretizing design elements in order to track the way they are perceived in a virtual environment (i. e. through eye tracking technology). Such an act necessitates deeper thinking by a designer – whether they design streetscapes, district form codes, plazas, or planted medians – to formulate a nuanced understanding of a "design element". Developing a mental map to formulate a design element, requires constructing different scales of elements in a meaningful hierarchy. This is analogous to a process we are already familiar with, such as when we arbitrarily discretize the continuous scale from site to region. Doing this at an elemental level that translates to a computational and perception-based context can reorient how the designer understands the parameters of perception (across scales). In this case, creating an eye tracking application can challenge those parameters (or inform their construction) by making the designer choose, *a priori*, boundaries of user perception or AOIs (whether as a collective or individual objects). For instance, if a designer or researcher designates a certain architectural or landscape detail (e. g. an alcove, seat wall, planting bed, etc.) as an "element" but the user of the space never fixates on it so narrowly, is it truly an element in itself or is it rather part of a larger one at the urban scale? In asking such questions, the landscape or urban designer will develop a deeper knowledge and higher sensitivity of elements that combine to create experience when encountering the space. Furthermore, they can take this more nuanced framework of discretized spatial elements and apply it broadly across project types and scales, sharing and reusing the code and data to bring the discipline closer to an ideal ontological foundation of digital landscapes (FERNBERG et al. 2021).

The second benefit we foresee is an extension of the first. In developing such an explicit, highly discretized framework of spatial elements and configurations, the digital landscape architect extends the benefits of gaze tracking in VR beyond more holistic critical thinking about design process into concrete controls for empirical assessments of design decisions. When eye tracking is done using 2D imagery or in real-world settings, statistical inferences about perception in the built environment can be highly susceptible to bias and confounding, either from lack of immersion and study imagery that is too parochial in the case of the former (DUPONT et al. 2016, HOLLANDER et al. 2020) or lack of controls for the volatile, multi-factorial conditions in real spaces, the case of the latter (JUNKER & NOLLEN 2018, RAJENDRAN & ODELEYE 2020, SIMPSON et al. 2019). Implementing eye tracking in the controlled conditions of a virtual environment helps to mitigate both situations by providing a more immersive, multi-sensory experience as well the power to choose which "things" are being measured in any given time or condition. In other words, the subjective measures often used to assess proposed designs and their future performance based on values and experience – such as surveys to gather feedback after a VR walkthrough in a laboratory study or a public open house for a potential project – can be enhanced and cross-examined with objective biometrics gathered outside of the participant's explicit awareness – in this case through gathering of gaze data – thus allowing for a unique and enlightening mixed methods approach to spatial design research and practice.

The first and most obvious challenge, we foresee researchers in landscape architecture and urban design experiencing, is the technological implementation of eye tracking in VR. Our research team has developed an implementation of the eye tracking software that will be made publicly available. While the software will alleviate a significant reliance on programming, it still requires expertise to integrate within the Unity gaming engine and setup as part of a programming interface. This will involve understanding how to organize and name game objects and attach the eye tracking plugin into the game scene. Once these are integrated, the setup with the HMD should go smoothly, as it is an already well-documented process.

The second most relevant challenge we anticipate is the skillset to analyse the complex data, as well as to make inferences about what these data mean. While eye tracking is an extraordinary tool on its own, its value is usually shown in its ability to elucidate variability in other data, such as showing whether poorly and well-remembered items in an environment received differing levels of the viewer's attention (LAVELLE et al. 2021). Similarly, eye-tracking data can be used within the context of landscape architecture and design to determine elements of a complex environment that pull the viewer's attention, but a certain level of expertise is still needed to examine reasons for the viewer to attend to or ignore specific design elements as well as the consequences for behaviour. Thus, one of the greater benefits mentioned above of a more grounded ontological understanding of design by way of choosing boundaries of perception also becomes one of VR-based eye tracking's greater obstacles to overcome.

4 Conclusions

As VR becomes a more common platform in design fields (both for practice and research), we must continuously assess the costs of implementation and benefits to identify potential uses. The extensive controls the platform affords and the integration with objective measurement medium, such as eye tracking, serve as a conduit to improve research and design process whilst enabling new kinds of performance measures. In this reflection paper we have highlighted two important benefits we have seen in our research: 1) a more holistic, discretized design process, and 2) greater empirical controls for design assessment. Further we offered two major challenges: 1) the technological implementation and 2) drawing associations between the data toward meaningful design decisions. As we look to further our research using eye tracking in 3D virtual environments, our next step is to answer the following research question: 1) to what extent can we empirically evaluate design scenarios using eye tracking and are these evaluations meaningful for informing design?

References

CARL, E., STEIN, A.T., LEVIHN-COON, A., POGUE, J. R., ROTHBAUM, B., EMMELKAMP, P., ASMUNDSON, G. J., CARLBRING, P. & POWERS, M. B. (2019), Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. Journal of Anxiety Disorders, 61, 27-36.

CARMONA, M. (2021), Public Places Urban Spaces: The Dimensions of Urban Design (3rd Ed.). Routledge. https://doi.org/10.4324/9781315158457.

DEMING, M. E. & SWAFFIELD, S. (2011), Landscape Architectural Research: Inquiry, Strategy, Design. John Wiley & Sons.

DUPONT, L., ANTROP, M. & VAN EETVELDE, V. (2014), Eye-tracking Analysis in Landscape Perception Research: Influence of Photograph Properties and Landscape Characteristics. Landscape Research, 39 (4), 417-432. https://doi.org/10.1080/01426397.2013.773966.

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. https://doi.org/10.1016/j.landurbplan.2015.12.007.

EMO, D. B., AL-SAYED, D. K. & VAROUDIS, D. T. (2016), Design, cognition & behaviour: Usability in the built environment. International Journal of Design Creativity and Innovation, 4 (2), 63-66. https://doi.org/10.1080/21650349.2016.1143080.

EWING, R. & CLEMENTE, O. (2013), Measuring Urban Design: Metrics for Livable Places. Island Press.

FERNBERG, P., STURLA, P. & CHAMBERLAIN, B. (2021), Pursuing an AI Ontology for Landscape Architecture. Journal of Digital Landscape Architecture, 6-2021. https://doi.org/10.14627/537705040.

FILOMENA, G., VERSTEGEN, J. A. & MANLEY, E. (2019), A computational approach to 'The Image of the City.' Cities, 89, 14-25. https://doi.org/10.1016/j.cities.2019.01.006.

HAMILTON, D., MCKECHNIE, J., EDGERTON, E. & WILSON, C. (2021), Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. Journal of Computers in Education, 8 (1), 1-32.

HERMUND, A., KLINT, L. & BUNDGÅRD, T. S. (2018), The Perception of Architectural Space in Reality, in Virtual Reality, and through Plan and Section Drawings. Computing for a Better Tomorrow, 36, 735-744.

HOLLANDER, J. B., SUSSMAN, A., PURDY LEVERING, A. & FOSTER-KARIM, C. (2020), Using Eye-Tracking to Understand Human Responses to Traditional Neighborhood Designs. Planning Practice & Research, 35 (5), 485-509. https://doi.org/10.1080/02697459.2020.1768332.

HOLMQVIST, K., NYSTRÖM, M., ANDERSSON, R., DEWHURST, R., JARODZKA, H. & WEIJER, J. VAN DE (2011), Eye Tracking: A comprehensive guide to methods and measures. OUP Oxford.

JUNKER, D. & NOLLEN, C. (2018), Mobile Eye Tracking in Landscape Architecture: Discovering a New Application for Research on Site. In: Landscape Architecture: The Sense of Places, Models and Applications. IntechOpen, 45-65.

KIM, N. & LEE, H. (2021), Assessing Consumer Attention and Arousal Using Eye-Tracking Technology in Virtual Retail Environment. Frontiers in Psychology, 12, 2861. https://doi.org/10.3389/fpsyg.2021.665658.

LAVELLE, M., ALONSO, D., LURIA, R. & DREW, T. (2021), Visual working memory load plays limited, to no role in encoding distractor objects during visual search. Visual Cognition, 29 (5), 288-309. https://doi.org/10.1080/13506285.2021.1914256.

LI, C.-L., AIVAR, M. P., KIT, D. M., TONG, M. H. & HAYHOE, M. M. (2016), Memory and visual search in naturalistic 2D and 3D environments. Journal of Vision, 16 (8), 9-9.

MORELLO, E. & RATTI, C. (2009), A digital image of the city: 3D isovists in Lynch's urban analysis. Environment and Planning B: Planning and Design, 36 (5), 837-853.

PEREZ-MARCOS, D. (2018), Virtual reality experiences, embodiment, videogames and their dimensions in neurorehabilitation. Journal of Neuroengineering and Rehabilitation, 15 (1), 1-8.

RAJENDRAN, L. P. & ODELEYE, N. D. (Eds.). (2020), Mediated Identities in the Futures of Place: Emerging Practices and Spatial Cultures. Springer International Publishing. https://doi.org/10.1007/978-3-030-06237-8.

REN, X. & KANG, J. (2015), Interactions between landscape elements and tranquility evaluation based on eye tracking experiments. The Journal of the Acoustical Society of America, 138 (5), 3019-3022. https://doi.org/10.1121/1.4934955.

SIMPSON, J., FREETH, M., SIMPSON, K. J. & THWAITES, K. (2019), Visual engagement with urban street edges: Insights using mobile eye-tracking. Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 12 (3), 259-278. https://doi.org/10.1080/17549175.2018.1552884.

SLATER, M. (2018), Immersion and the illusion of presence in virtual reality. British Journal of Psychology (London, England: 1953), 109 (3), 431-433. https://doi.org/10.1111/bjop.12305.

WALLER, D. (2005), The WALKABOUT: Using virtual environments to assess large-scale spatial abilities. Computers in Human Behavior, 21 (2), 243-253. https://doi.org/10.1016/j.chb.2004.02.022.

XIANGMIN, G., WEIQIANG, C., TIANTIAN, L. & SHUMENG, H. (2021), Research on dynamic visual attraction evaluation method of commercial street based on eye movement perception. Journal of Asian Architecture and Building Engineering, 2021, 1-13. https://doi.org/10.1080/13467581.2021.1944872.

ZHANG, F., ZHOU, B., LIU, L., LIU, Y., FUNG, H. H., LIN, H. & RATTI, C. (2018), Measuring human perceptions of a large-scale urban region using machine learning. Landscape and Urban Planning, 180, 148-160. https://doi.org/10.1016/j.landurbplan.2018.08.020.

ZHANG, L.-M., ZHANG, R.-X., JENG, T.-S. & ZENG, Z.-Y. (2019), Cityscape protection using VR and eye tracking technology. Journal of Visual Communication and Image Representation, 64, 102639. https://doi.org/10.1016/j.jvcir.2019.102639.