

**How veridical are different modalities of digital representation?
The effect of presentation modality on physiological response**

Author

Gero, John S, Shields, Jennifer, Yu, Rongrong

Published

2016

Conference Title

PROCEEDINGS OF THE 21ST INTERNATIONAL CONFERENCE ON COMPUTER-AIDED
ARCHITECTURAL DESIGN RESEARCH IN ASIA (CAADRIA 2016)

Version

Version of Record (VoR)

Rights statement

© 2016 CAADRIA. The attached file is reproduced here in accordance with the copyright policy of the publisher. For information about this conference please refer to the conference's website or contact the author(s).

Downloaded from

<http://hdl.handle.net/10072/340510>

Link to published version

<https://www.unimelb.edu.au/>

Griffith Research Online

<https://research-repository.griffith.edu.au>

HOW VERIDICAL ARE DIFFERENT MODALITIES OF DIGITAL REPRESENTATION?

The effect of presentation modality on physiological response

JOHN S. GERO¹, JENNIFER SHIELDS² and RONGRONG YU³

¹*Krasnow Institute for Advanced Study, George Mason University, Fairfax, USA and University of North Carolina at Charlotte, USA
john@johngero.com*

²*California Polytechnic State University, San Luis Obispo, USA
jeshield@calpoly.edu*

³*Northeast Forestry University, Harbin, China
rongrong.yu@uon.edu.au*

Abstract. This paper presents the methodology and results of a pilot study comparing the eye movements of architecture students when looking at different modalities of digital spatial representation. The two participant groups consisted of third and fourth year undergraduate architecture students from the University of North Carolina at Charlotte, USA and from Harbin Institute of Technology, Harbin, China. The two modalities studied were a computer-generated line drawn perspective and a digitized photograph of the same architectural space. The results of this study show that student designers' physiological response varies with different representation modalities.

Keywords. Eye movement; spatial representation; modalities of spatial representation; designers' physiological response.

1. Introduction

Architects design with varying modalities of representation; these include sketches, drawings, physical models, and digital models to assist in the development of design concepts. Visual representations during the design process assist designers with their concept development (Schön, 1992). With the increasing application of computers in architectural design, the effect of digital representations on designers during the design process is not understood. With this pilot project we aim to produce evidence-based results about de-

signers' physiological responses as they are exposed to different modalities of spatial representation. This study is part of the project which aims to relate designers' responses in the embodied experience of the built space with the responses evoked when looking at digital models.

This pilot study utilized eye tracking to collect physiological data, comparing responses to different spatial representations. Building on the research paradigm of Scene Perception (Rensink, 2000), researchers have been interested in how designers look at scenes. Experiments typically present the scene on a computer screen, allowing for a high level of control over variables of the scene. All modalities of spatial representation are abstractions of physical space and are therefore subject to interpretation. From the analysis of the relationship between designers' eye tracking data and different modalities of representation, designers' responses to spatial representation will be presented and compared.

The remainder of the paper presents a brief background to computer representation, spatial perception and eye tracking. This is followed by the research aim, significance and research design. From an analysis of the eye-tracking data, results are presented and discussed.

2. Background

2.1. DIGITAL REPRESENTATION

Design representation is a core issue in most design domains including product design, architecture and engineering. With the development of computational modelling, digital design representation becomes possible during the design process, which assists designers in both off-loading cognition and providing the possibility to interact with their external representations (Schön and Wiggins, 1992). For example, BIM technology enables 3D (model check, design view, enhanced reality) and 4D visualization (plus time) (Eastman, 2008). Virtual reality (VR) can provide realistic virtual environment which enable designers navigation possibilities (Wang and Dunston, 2013); augmented reality (AR) can enhance the user's perception by complementing the real world with 3D virtual objects in the same space (Morrison et al., 2011). These digital representation tools allow designers to more readily explore design ideas and assist with the concept development of their designs. However, there is a lack of understanding of the physiological effect of digital representations on designers.

2.2. SPATIAL PERCEPTION

Oliva et al (2011) studied the representation of the shape of visual space. In their study they critique the isovist theory of spatial perception for leaving out other characteristics of space such as texture, material and colour that contribute to the human perception of space. They also propose a model called the 'spatial-envelope representation', which describes qualitatively the character and mood of a physical or pictorial space, represented by its boundaries (e.g., walls, floor, ceiling, and lighting) stripped of movable elements (e.g., objects and furnishing). Then they proposed a formal, computational approach to the capture of the shape of space as it would be perceived from an observer's vantage point (Oliva and Torralba, 2001; Oliva and Torralba, 2006; Oliva et al, 2011). The collection of properties describing a space in view is referred to as the spatial-envelope representation. This distinction in the characteristics that contribute to the human perception of space pertains to our research into representations of space that include or exclude some of those characteristics.

2.3. EYE TRACKING

Early research on eye movement can be traced back to Buswell (1935) who focused on the aesthetic impact of photographs of artwork, patterns and sculpture, particularly the layout patterns of advertisements. Kaufman and Richard (1969) measured eye fixation times in several pre-defined parts of figures, the results show that the center of gravity is an attractor as are the edges and corners. Gould and Peeples (1970) suggest that a subject's interpretation of a figure does not affect eye movements, which means that only "physical attributes" have influence on the eye movements. Torralba et al (2006) proposed visual attentional guidance through an experimental search task. Results of their study suggest that the context information plays an important role in object detection and observation. They also suggest some parts of the scene attract more attention than others.

While the relationship between eye movement and perception of artworks has been investigated, there has been very little study on the role of eye movement in the perception of three-dimensional architectural space. One of the few studies on this topic was conducted by Weber, Choi, and Stark (2002) in which they collected eye tracking data as participants were asked to look at three-dimensional models, or photographs of models, of architectural space. These models were constructed to collect data on the perception of the following architectural issues: empty space; symmetry vs. asymmetry; left and right reversed; obliquely-oriented elements; vista; and foreground. The research focused on comparing different arrangements of objects within

a space, rather than different methods of representing the same spatial configuration. Their results showed that, with no priori model in a figure, the attention would fixate at the center; while the foreground was common for initial fixations, the eye did not typically scan the edges of interior space or rectilinearly-oriented contours; the objects on the left attract more attentions than the ones on the right. This confirmed the results by Arnheim (1985). The study also concluded that fixations did not vary significantly when viewing the three-dimensional model compared with a photograph of the model, with the exception of the foreground, which attracted greater attention in the 3D model.

3. Aim

In the present study, we studied the eye movements of architecture students in the US and China in different modalities of digital spatial representation. The aim of this research is to explore the physical responses (eye movement) to spatial representation of architecture students when they view spatial scenes. To determine whether computer-aided representation of a set of spaces is veridical with a digitized photograph representation of the same set of spaces, measurements of physiological response – eye movement is reported here – to both a computer generated 3D model and a digitized photograph of the same set of spaces are undertaken.

Through the collection of eye tracking data, our aim was to test the preliminary hypothesis that physiological responses vary with representational modality.

4. Significance

As a starting point, our preliminary investigation focuses on the relationship between the perceptions of space in different representational modalities for subjects trained in architectural design. A future larger-scale study aims to identify the modalities of representation that most closely correlate with the embodied architectural experience to predict how spaces will affect the occupant's physiological state. The potential to predict physiological responses to a constructed space before that space exists would be an invaluable design tool. This could be achieved by collecting and analysing data for the anticipated user group's physiological responses to the most veridical representations of that space.

5. Research design

To achieve the preliminary hypothesis we measure physiological response to representations, in particular, eye movement. In this experiment, 22 third and

fourth year architecture students at UNC-Charlotte (UNCC) and 30 third and fourth year architecture students at Harbin Institute of Technology (HIT) in China participated in this pilot study.

During the experiment participants were asked to complete demographic questions regarding their gender, age and first language. They were required to look at two images shown on a screen (Fig.1): the first (R1) is a computer generated perspective drawing, the second (R0) is a digitized photograph of the same space when built. When they looked at the images, their eye-tracking data were recorded by an eye-tracking system (Gazepoint in USA and Tobii studio in China). Each of the images was shown for 20 seconds with a few seconds for recalibration in between them. Half of participants were first shown R1 then R0. The other half was first shown R0 then R1. After each image session there was a questionnaire session. Data collected included eye fixations and saccades.



*Figure 1. Left: Computer generated perspective drawing (labelled R1 in the experiment).
Right: Digitized photograph of the same space (labelled R0)*

We identified 7 Areas of Interest (AOIs) in the visual scene presented (Fig. 2). Each AOI defines an area that we wanted to gather data about. AOIs defined the three doorways, the two wall surfaces between them, the terminus of the corridor and the ceiling.

6. Results

Four metrics of eye movement data were captured during the experiment: Time to First View (secs), Time Viewed (%), Fixations and Average Revisits. Time to First View measures how long it takes before a test participant

fixates on an active AOI for the first time. Time Viewed (%) is the percentage of time viewed within an active AOI of the total viewing time – 20 seconds in this experiment. Fixations measure the number of times the participant fixates on an AOI. The Average Revisits measures the number of visits within an active AOI. A visit is defined as the time interval between the first fixation on the active AOI and the end of the last fixation within the same active AOI where there have been no fixations outside the AOI.

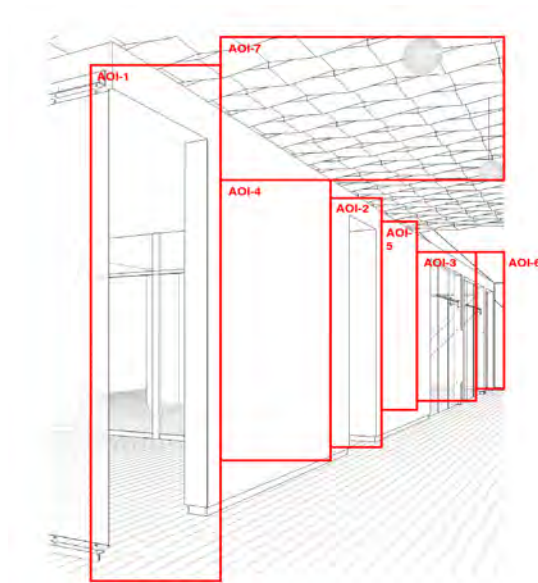


Figure 2. 7 Areas of Interest labelled AOI 1 through AOI 7 in R1, they occupy the same positions in R0.

There were no significant differences in the results based on the order of viewing of the images.

Table 1 shows the averages of the four metrics for the data of the 7 AOIs in Image R1. Table 2 shows the average for each of the four metrics of eye movement data of the 7 AOIs in Image R0.

We next determine whether there are statistically significant differences between these two representation modalities. Table 3 presents the t-test results of comparing Image R1 to Image R0.

VERIDICALITY OF MODALITIES OF DIGITAL REPRESENTATION? 867

TABLE 1. The average of eye-tracking data of 7 AOIs in Image R1.

AOI	1	2	3	4	5	6	7
UNCC students							
Time to First View (secs)	2.84	4.39	2.04	3.50	2.98	3.70	1.14
Time Viewed (%)	14.68	3.61	9.90	4.07	2.66	3.56	12.74
Fixations	12.58	5.18	10.18	6.09	5.55	4.64	13.60
Average Revisits	6.00	4.30	7.73	5.20	5.30	4.80	7.30
HIT students							
Time to First View (secs)	1.16	3.18	0.95	1.44	2.79	5.93	2.50
Time Viewed (%)	24.06	7.05	21.43	6.04	2.10	6.62	12.10
Fixations	19.61	5.87	11.83	6.30	1.83	4.97	9.70
Average Revisits	6.78	4.03	7.10	4.40	1.63	2.97	4.30

TABLE 2. The average of eye-tracking data of 7 AOIs in Image R0.

AOI	1	2	3	4	5	6	7
UNCC students							
Time to First View (secs)	0.86	3.24	2.31	1.71	4.13	7.83	2.16
Time Viewed (%)	22.16	3.00	10.12	4.06	2.79	4.51	8.49
Fixations	19.71	4.78	9.52	5.52	5.13	4.70	8.48
Revisits	8.13	4.10	6.70	3.95	4.95	4.00	4.46
HIT students							i
Time to First View (secs)	0.53	4.63	2.34	0.76	3.70	7.27	4.86
Time Viewed (%)	28.51	5.82	14.48	6.58	3.08	4.21	8.99
Fixations	28.70	5.93	12.17	8.47	3.83	3.97	8.77
Revisits	8.73	3.80	6.77	5.27	3.30	2.83	3.93

Table 3. t-test of Image R1 compared with Image R0.

AOI	p values						
	1	2	3	4	5	6	7
UNCC Students							
Time to First View (secs)	0.001*	0.558	0.661	0.188	0.412	0.049*	0.224
Time Viewed (%)	0.031*	0.635	0.760	0.888	0.803	0.388	0.234
Fixations	0.042*	0.861	0.895	0.814	0.900	0.799	0.123

Revisits	0.062	1.000	0.621	0.393	0.832	0.478	0.165
HIT Students							
Time to First View (secs)	0.001*	0.204	0.030*	0.633	0.400	0.353	0.029*
Time Viewed (%)	0.028*	0.267	0.009*	0.633	0.154	0.067	0.026*
Fixations	0.003*	0.960	0.873	0.633	0.015*	0.359	0.631
Revisits	0.002*	0.751	0.667	0.633	0.007*	0.820	0.591

* $p < 0.05$

7. Discussion and conclusions

The results of eye-tracking measurements are discussed initially in terms of the architecture students' behavior when viewing each image. This is followed by a discussion comparing the results of the eye-tracking measurements for both images to determine whether different modalities of digital representation produce different physiological behavior in the viewer.

The results in Tables 1 and 2 show how students allocated their cognitive effort. For example, from Table 1, we can see that the participants from both UNCC and HIT have the longest time viewed in AOI 1, largest average revisits in AOI 3 and the most fixations on AOI 7, AOI 1 and AOI 3. The least cognitive effort is allocated in AOI 2, AOI 5 and AOI 6.

There are two possible reasons for the long time viewed and fixation in AOI 1. Firstly, participants were attracted by the relatively more complex spaces in a spatial figure - AOI 1 is an open door through which they could see a glass wall inside the room, as well as some spaces behind the glass wall. When a designer is trying to understand the space, he/she will be attracted by this kind of complexity. Secondly, AOI 1 is located on the left hand side of the image, and according to study of Weber et al. (2002) and Arnheim (1985) the left side of a composition attracts more attention than the right. AOI 7 is the ceiling, which contains some complex structure, therefore it attracts the attention of the participants. The most revisited is AOI 3. This may be because compared to AOI 1 it is relatively small and distant so that participants need to revisit to try to understand the space. The least attention is directed at AOIs 2, 5, and 6 possibly because they are relatively simple areas.

The results in Table 3 indicate differences in three of the four eye movement measurements. Time to first view, time viewed and fixations were all statistically significantly different when viewing the computer-generated line drawing perspective compared to viewing the digital colour photograph for Area of Interest 1, which is the first AOI to be viewed in the digital photo-

graph but is only the third AOI to be viewed in the line drawing perspective. Interestingly, AOI-1 had the highest percentage of time viewed for both modalities, whilst there was a lack of congruence for the other three measures.

The results from this study additionally imply the following:

- Architectural students put more cognitive effort on the relative complex spatial representation areas, as well as the structural complex parts.
The complex spatial representation areas such as AOI 1 naturally attracts more attention probably because it needs more exploration to understand the space. Certain features are more likely to attract eye focus, for example, in a face figure the eye focus is usually on “eye” and “mouth” (Gould and Peeples, 1970). In architecture, past research suggest that architects pay more attention to the spatial arrangement of various architectural elements (Weber et al., 2002). The more complex spatial area such as AOI 1 contain multiple architectural elements: the extensive rooms behind the opening and the glass door. Also the complex structural parts in the ceiling attracted more attention.
- The left-hand side parts in an image are likely to attract more attention.
- Research on eye movement of static images has shown that the left-hand side parts attract more attention than the right (Arnheim, 1985; Weber et al., 2002). The present research confirms this claim that human eyes tend to “rest” at the left hand side. The possible reason may due to the visual pattern recognition and memory system of the brain. Further exploration of this phenomenon in spatial representations is needed.
- The edges of the interior space were not frequently scanned.
- The most fixated areas are the complex spatial areas, rather than the edges or corners of the interior space. This result agrees with Weber et al. (2002)’s research on eye-tracking of interior space. However, in Kaufman and Richard (1969)’s study, they state that icons such as corners and edges are more likely to attract attention, while the current study does not confirm this claim.
- The center of gravity is not always the focus of eye movement.
- In the early eye tracking research, researchers found that the eye fixation is always focused on the center of gravity of a figure (Kaufman and Richard, 1969), however in the current study, the center of the image is not the focus on eye fixation, instead it is the relatively complex space. This indicates that the semantics of the image plays a role.

This paper has presented the results of a pilot study comparing the physiological response of eye movement of architecture students when looking at different digital modalities of spatial representation. Results from the analysis of eye-tracking data suggest that designers’ physiological responses vary with the changes of digital representation. In order to make a more comprehensive conclusion, a larger experiment is needed in a future study. Future

studies would include a broader demographic (comparing trained designers with non-designers) and a wider range of representational modalities.

References

- Arnheim, R. 1985. *The power of the center*, Berkeley, University of California Press.
- Buswell 1935. *How people look at pictures: A study of the psychology of perception in art*, Chicago, University of Chicago Press.
- Eastman, C. M. 2008. *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*, Wiley.
- Gould, D. & Peeples, R.: 1970, Eye movements during visual search and discrimination of meaningless symbol and object patterns, *Journal of Experimental Psychology*, **85**, 51–55.
- Kaufman, L. & Richard, W.: 1969, Spontaneous fixation tendencies for visual forms, *Perception and Psychophysics*, **5**, 85–88.
- Morrison, A., Mulloni, A., Lemmelä, S., Oulasvirta, A., Jacucci, G., Peltonen, P., Schmalstieg, D. & Regenbrecht, H.: 2011, Collaborative use of mobile augmented reality with paper maps, *Computers & Graphics*, **35**(4), 789–799.
- Oliva, A., Park, S. & Konkle, T.: 2011, Representing, perceiving and remembering the shape of visual space. , in: Harris, L. R. & Jenkin, M. (ed), *Vision in 3D Environments*, Cambridge University Press, Cambridge. Pp. 308–339.
- Oliva, A. & Torralba, A.: 2001, Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope, *International Journal of Computer Vision*, **42**(3), 145–175.
- Oliva, A. & Torralba, A.: 2006, Building the gist of a scene: the role of global image features in recognition, *Progress in Brain Research*, **155**, 23–36.
- Rensink, R. A.: 2000, Scene Perception, in A. Kazdin (ed), *Encyclopedia of Psychology*, Oxford University Press, New York. 151–155.
- Schön, D. & Wiggins, G.: 1992, Kinds of seeing and their functions in designing, *Design Studies*, **13**(2), 135–156.
- Schön, D. A.: 1992, Designing as reflective conversation with the materials of a design situation, *Knowledge-Based Systems*, **5** (1), 3–14.
- Torralba, A., Oliva, A., Castelano, M., S & 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**(4), 766–786.
- Wang, X. & Dunston, P.: 2013, Tangible mixed reality for remote design review: a study understanding user perception and acceptance, *Visualization in Engineering*, **1**(1), 1–15.
- Weber, R., Choi, Y. & Stark, L.: 2002, The Impact of Formal Properties on Eye Movement During the Perception of Architecture, *Journal of Architectural Planning and Research*, **19**(1), 57–68.