Evaluating the veridicality of two-dimensional representations of three-dimensional architectural space through physiological response.

Jennifer A. E. Shields
California Polytechnic State University, San Luis Obispo, CA, USA

John. S. Gero
Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA, USA and University of North Carolina at Charlotte, NC, USA

Rongrong Yu
The University of Newcastle, Newcastle, Australia

ABSTRACT: This paper presents research on the perception of three-dimensional space in two-dimensional media, and the methodology and results of a pilot study comparing the eye movements of architecture students when looking at different modalities of spatial representation. The results of this study show that student designers' physiological responses vary with different representation modalities.

1 INTRODUCTION

Architects design with varying modalities of representation, including sketches, drawings, physical models, and digital models. These representations may serve to assist architects in the development of design concepts (Schön, 1992), or to communicate design intentions to the client and/or future users. The built work that results from a design process is a new configuration of spaces and forms that humans respond to both emotionally and physiologically. But how does the built work compare with its earlier representations? All modalities of spatial representation are abstractions of physical space and are therefore subject to interpretation. Since graphical representations are the most common form of representation of space during designing it is important to know what the implications for the designer are in representing space in different graphical modalities. The research outlined in this paper is part of a project which aims to relate physiological responses in the embodied experience of the built space with the responses evoked when looking at architectural representations. This research could have substantial implications for our design process as architects and our pedagogical methods as architectural educators, in our challenge to bridge the divide between concept and built work.

Little empirical research has been done on the human physiological response to space and spatial representations. This project is a cross-disciplinary investigation involving researchers in architecture, computer science and psychology. Our research questions include: How veridical are these spatial representations? Do physiological responses to modalities across a participant group demonstrate consistencies? Are there cultural differences in physiological responses? Can physiological responses to space as represented in various modalities of architectural representation be correlated with physiological responses to physically inhabited space?

For the purposes of this pilot study, we narrowed our focus to pose the previously unexplored question: with the increasing application of computers in architectural design, what is the effect of different modes of digital representation on designers? With this pilot project we aim to produce evidence-based results about designers' physiological responses as they are exposed to two different modalities of spatial representation.

This pilot study utilized eye tracking to collect physiological data, comparing responses to different spatial representations. Through the collection of eye tracking data, our aim was to test our preliminary hypothesis: that physiological responses vary with representational modality. Building on the research paradigm of Scene Perception researchers (Rensink, 2000) 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 (Holmqvist 2011). Our pilot study follows this methodology. From the analysis of the relationship between designers' eye tracking data and different modalities of representation – perspectival line drawing and photograph - designers' responses to spatial representation will be presented and compared.

The remainder of the paper presents a brief background to architectural 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 analysed, and broader impact discussed.

2 BACKGROUND

2.1 *Architectural Representation*

Architectural drawings are understood as referential and non-autonomous; they are representations of architectural space and form. Plans, sections, and perspectival views typically follow discipline-specific conventions for representation. These drawings may specify the building, or they may depict it: defined as notational or imaginative drawings by Sonit Bafna (2008). This distinction accounts for the drawing-as-instruction, as well as the drawing-as-provocation. According to Bafna, "Richard Wollheim has argued that depictive seeing activates a special mechanism that he termed 'seeing-in'; faced with such pictures one cannot help but see figures within them despite knowing full well that they are just marks on the surface." Wollheim claims that we simultaneously perceive the drawing as an object, and the objects/spaces depicted within the drawing, with a "two-fold sense of perception." (Bafna 2008) While the intent of the imaginative drawing is not to evoke the experience of the building depicted, it is meant to provoke a "mode of attention," allowing for an engagement with the representation that allows for multiple readings (Bafna 2008).

With photography, however, there tends to be an assumption that it is an unbiased presentation of the space. Yet understanding the content and spatial implications of a photograph is a learned ability, and the photograph is, like a drawing, an abstraction of the architectural space. While studies comparing different cultures have shown no difference in perceptual organization, "the interpretation of perceptual material may differ considerably between two cultural groups." (Weber 1995) For example, African aborigines could not recognize anything in photographs they were shown. Two-dimensional representations of three-dimensional space requires interpretation which "relies on acquired visual conventions that may be as arbitrary as linguistic conventions" (Weber 1995). Photographs are authored and abstracted depictions of architectural space in the same manner as drawings.

2.2 *Spatial Perception*

The ambiguity inherent in any architectural representation of space is a characteristic of spatial perception as well. Rudolf Arnheim emphasized the relative nature of perception, describing perceptual experience as being an aggregate of natural and manmade objects. This aggregation forms a "spatial framework" that causes visual percepts to "contain more than what is given in the physical stimulus pattern" (Arnheim 1977). In other words, our experience of an architectural space is influenced as much by the relationships and tension (or lack of tension) between architectural elements as by those elements themselves.

Michael Benedikt's isovist theory helps designers to evaluate visibility within or from a space from the vantage point of an occupant, considering the physically bound space in comparison with what can be visually perceived. An isovist is the volume of space visible from a given point (Benedikt 1979). Oliva et al. (2011) also 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 color 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 by the authors 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*

Eye tracking data includes fixations and saccades, based on movement of the fovea. The fovea represents less than 2 degrees of the visual field, so the eye must move, or *foveate*, in order to take in detailed information (for example, in areas of an image with complexity). When the eye stops to take in information, this is called a *fixation*. The rapid movement of the eye from one fixation to another is called a *saccade*. During saccades, the eye is effectively blind: visual information obtained by the eye occurs during fixations, thus making fixations the most valuable data collected through eye tracking (Holmqvist 2011).

As early as the 1930s, psychologists were interested in how the eye moves and fixates when focused on a work of art. Guy Buswell (1935), a psychologist who invented the first non-intrusive eye tracker, found that the eyes do not follow edges, but tend to scan and focus on central concave areas. More recent studies confirm that fixations are likely to occur in concave and enclosed areas of a figure, rather than in what is perceived as negative space (Weber 2005).

Yarbus (1967) was another pioneer in eye tracking research, who found that the eyes move in different paths when looking at an image depending on the task the observer was asked to complete, Figure 1.
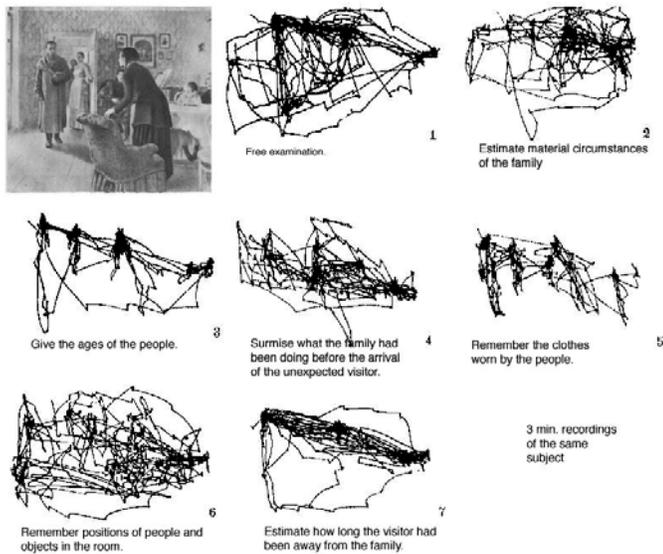

Figure 1. Eye tracking experiment by Alfred Yarbus in which participants were given a series of tasks while looking at Ilya Repin's painting 'An Unexpected Visitor' (Yarbus 1967).

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 fix-ate 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 (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.

While these studies did not consider cultural background in their data analysis, a 2005 study by Chua et al. tested the hypothesis that those from Western cultures (in this case, American participants) have a tendency to look at focal objects, while those from Eastern cultures (in this case, Chinese participants) tend to look more at context and background. This hypothesis was confirmed by their study. The question of possible cultural differences in eye movement also arose in our study.

3 PILOT STUDY

3.1 *Research Design*

To test the preliminary hypothesis that physiological responses vary with representational modality, we set up an experiment that measured eye movement in response to spatial representations. 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, Figure 2.

Figure 2. Eye tracking experiment setup by authors. This photo shows the setup at HIT in China; a similar setup was used in Charlotte.
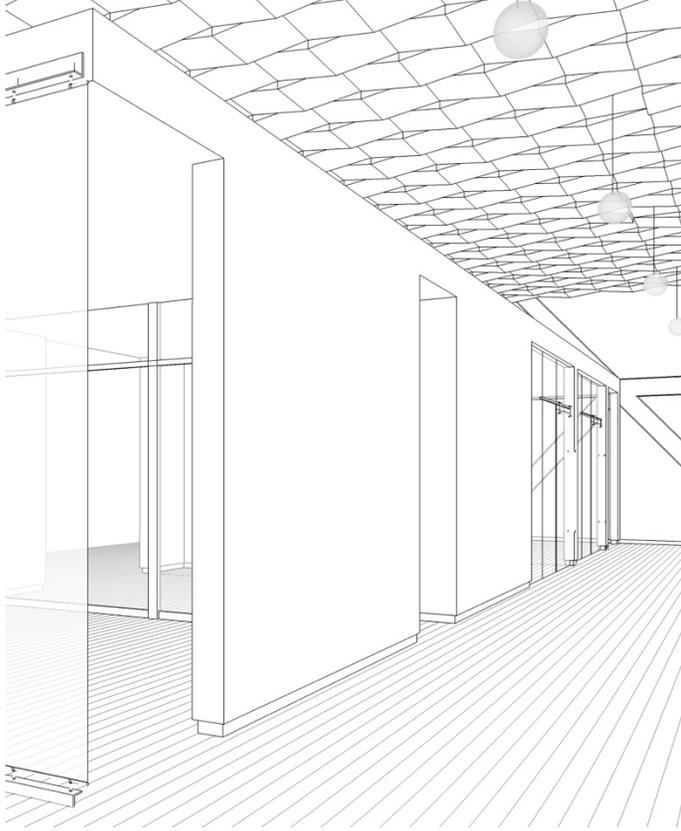
Figure 3. Computer generated perspective drawing (labelled R1 in the experiment).

Figure 4. Digitized photograph of the same space (labelled R0).

The first image, labeled R1, Figure 3, is a computer generated perspective drawing. The second image, labeled R0, Figure 4, 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.

3.2 *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.

We identified 7 Areas of Interest (AOIs) in the visual scene presented, Figure 5. 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.
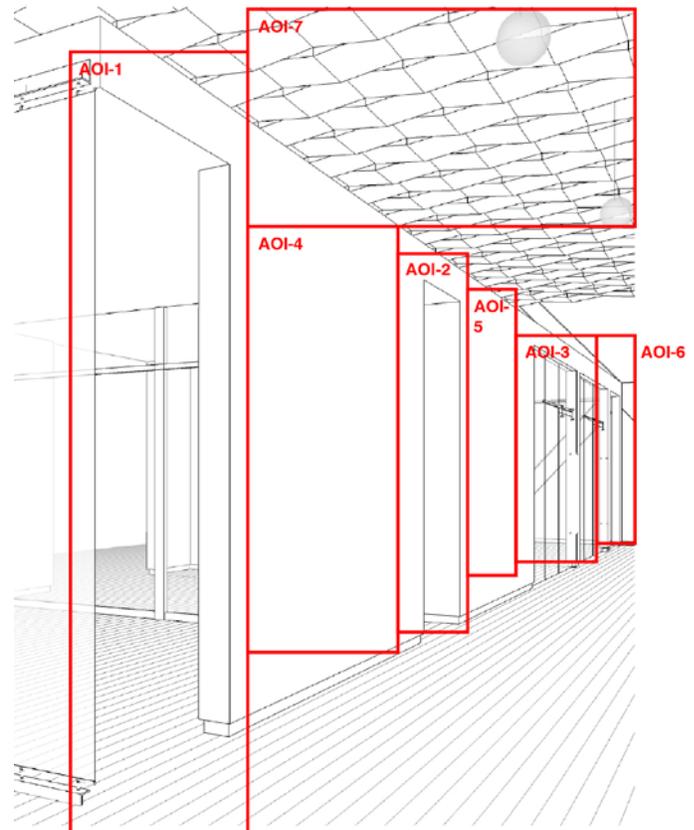

Figure 5. Seven Areas of Interest labeled AOI 1 through AOI 7 in R1; they occupy the same positions in R0.
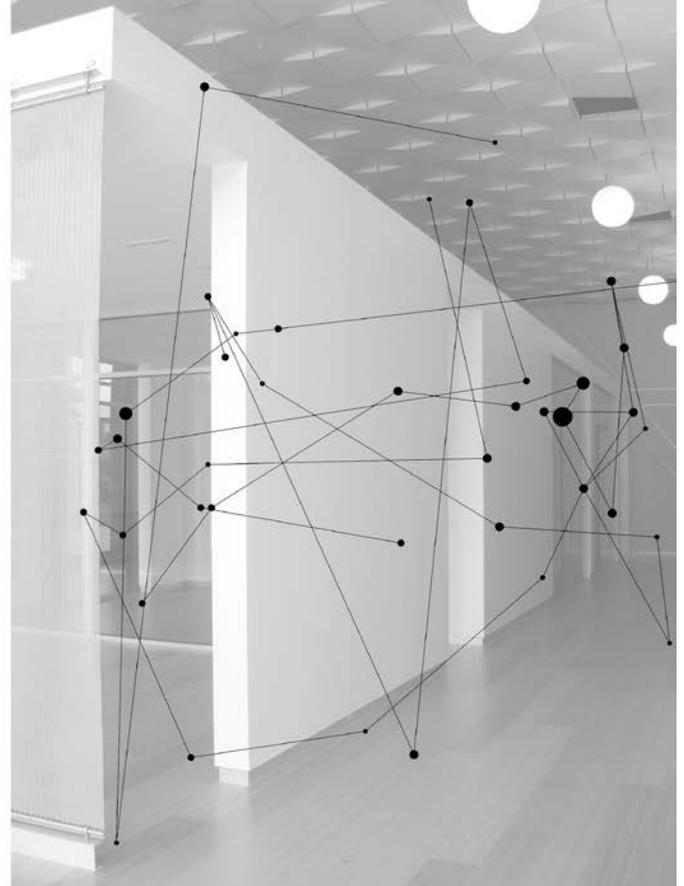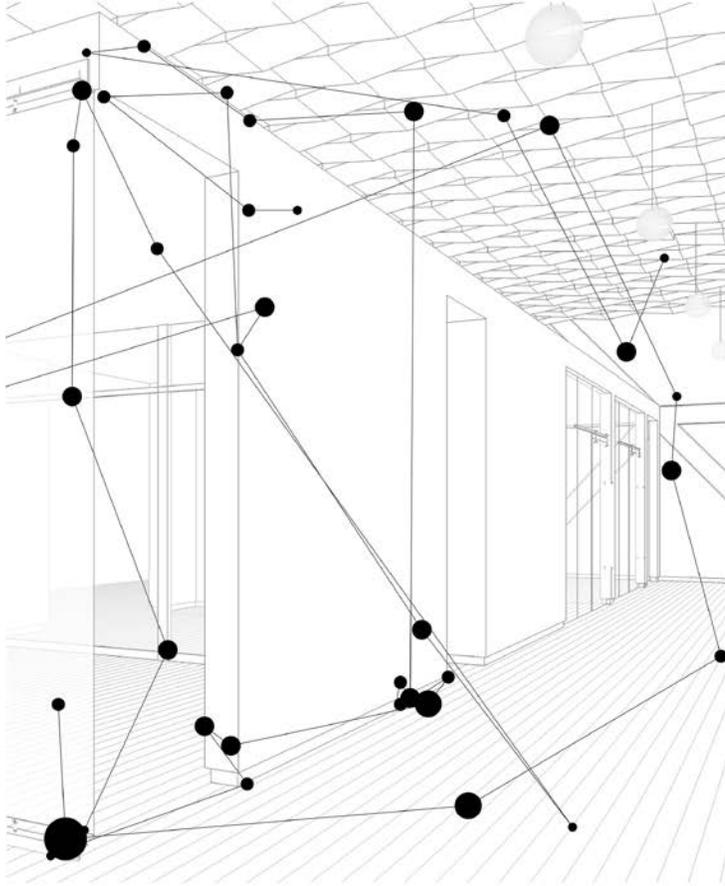
Figure 6. Example of scan path and fixations on R1 and R0.

Figure 6 gives a graphic example of the type of data collected: scan path and fixations for one participant are mapped to the images R1 and R0.

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. Statistically significant differences are marked with an asterisk.

4 DISCUSSION AND CONCLUSIONS

4.1 *Analysis of Results*

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 longer time spent viewing and fixating on AOI 1. Firstly, participants may have been 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. Also, fixations tend to occur within a contained figure, rather than in the perceived negative space. While the white architectural volume sits as a figure in the larger space depicted, its two-dimensional representation causes the white wall to be perceived as negative space. (The contour becomes part of the figure it encloses, rendering the ground boundless.) 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 statis-

tically significantly different when viewing the computer-generated line drawing perspective compared to viewing the digital color photograph for AOI 1, which is the first AOI to be viewed in the digital photograph 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 expend more cognitive effort on the relative complex spatial representation areas, as well as the structurally complex parts.

The complex spatial representation areas such as AOI 1 naturally attract 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.

- Cultural background of the student did not produce different results.

No significant difference was noted between the focal attention paid to AOIs by the American students as compared with the Chinese students. A previous study by Chua et al. (2005) confirmed the hypothesis that Americans fixated more on focal objects while Chinese participants spent more time looking at the context/background. This distinction was not noted in our pilot study.

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.

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

4.2 Next Steps

Our goal is to integrate and expand on research conducted in a number of fields, including architecture, computer science, and psychology. Researchers in these fields have investigated the processes by which we perceive and understand space, by collecting and analyzing physiological data. We intend to further our research by conducting a larger-scale study with more participants, in which the sequence of modalities can be varied, and in which we can test the participants' physiological responses to real space in addition to representations.

The two key sources of relevant data for our next study are eye trackers and EEG. As a counterpoint to eye tracking as used in our pilot study, we are interested in additional physiological responses evoked by a space and its representation, which can be measured by EEG. Researchers have employed these tools to track brain activity and stimulation when viewing an image, as well as the emotional responses evoked by real-world spatial experiences.

Our methodology for the next experiment will focus on the emotional responses to these images as indicated by EEG data, as an additional layer to the eye tracking data. The Emotiv EPOC is a portable EEG headset that collects brain wave data to identify levels of cross-modal sensory processing, alertness/engagement, reflection, and engagement of memory in gamma, beta, alpha and theta waves respectively. As a portable device, we will be able to use it in subsequent studies in which the participants' physiological responses are measured in an embodied experience of a built space.

An anticipated outcome of a future large-scale study would be the identification of modalities of representation that most closely correlate with the embodied architectural experience to predict how spaces will affect the occupant's physiological state. As we improve our understanding of the relationship between our representations of architecture space in the design process and how that space as a built work will ultimately be experienced, architectural educators can work with students to represent their design work in a more authentic way.

REFERENCES

Arnheim, R. 1977. *The Dynamics of Architectural Form.* Berkeley: University of California Press.

Arnheim, R. 1985. *The Power of the Center.* Berkeley: University of California Press.

Bafna, S. 2008. How Architectural Drawings Work - and What That Implies for the Role of Representation in Architecture. *Journal of Architecture*, 13.5: 535-564.

Benedikt, M. 1979. To take hold of space: isovists and isovist fields. *Environment and Planning*, 6: 47-65.

Buswell, G. 1935. *How People Look at Pictures: A study of the psychology of perception in art.* Chicago: University of Chicago Press.

Chua, H. F., Boland, J. E., and Nisbett. R. E. 2005. Cultural Variation in Eye Movements During Scene Perception. *Proceedings of the National Academy of Sciences of the United States of America*, 102.35: 12629-12633.

Gero, JS, Shields, J and Yu, R 2016. How veridical are different modalities of digital representation? *CAADRIA 2016* (to appear).

Gibson, J.J. 1986. *The Ecological Approach to Visual Perception*. New York: Psychology Press.

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.

Holmqvist, K. et al. 2015. *Eye Tracking: A comprehensive guide to methods and measures*. New York: Oxford University Press.

Kaufman, L. & Richard, W. 1969. Spontaneous fixation tendencies for visual forms. *Perception and Psychophysics,* **5**, 85-88.

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: Cambridge University Press. 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:* KAZDIN, A. (ed), *Encyclopedia of Psychology,* New York: Oxford University Press. pp. 151-155.

Schön, D. A. 1992. Designing as reflective conversation with the materials of a design situation, *Knowledge-Based Systems,* **5** (1), 3-14.

Segall, M.H. et al. 1966. *The Influence of Culture on Perception*. Indianapolis: Bobbs-Merrill.

Torralba, A., Oliva, A., Castelhano, 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.

Weber, Ralf. 1995. *On the Aesthetics of Architecture: A Psychological Approach to the Structure and the Order of Perceived Architectural Space*. Aldershot: Avebury.

Yarbus, A. and Haign, B. 1973. *Eye Movements and Vision*. New York: Plenum Press.