

Analysis of Signage Using Eye-Tracking Technology

Ming Tang

Associate Professor
School of Architecture and Interior Design,
College of Design, Architecture,
Art, and Planning
University of Cincinnati

tangmg@ucmail.uc.edu

SIGNAGE AS COMMUNICATION

Signs, whether freestanding, building-mounted or interior, exist in a variety of forms and manifestations. They may include a combination of letters, words, graphics, and symbols composed in a variety of colors, materials, textures, and lighting. Whatever their composition, signs can provide essential visual communication for wayfinding, commerce, and public dialogue and expression. Those interested in how the design, placement, and potential regulation of signs may impact their effectiveness will appreciate that signs are viewed in the context of their visual environment, with their impact on human behavior reflecting a complex array of experiential, psychological, and physiological factors.

To have their intended impact, signs need to capture the subconscious visual attention of viewers. Vision science has established that in humans this pre-attentive viewing or visual attention is innately drawn to visual elements that contains lines and edges, red/green and blue/yellow contrasts, intensity against context, and facial features within the visual environment; these visual elements are considered triggers of the conscious viewing required for response (Moore, 2011). Through the perception of form, spatial configuration, light, and material, the contextual built and natural environments influence how viewers experience signage. As with other elements of the built environment, signage is best understood to be experienced in its full context and has its impact through a combination of sensation and perception, learning and memory, decision making, emotion and effect, and movement (Eberhard, 2009). Signage design, to be effective, must appreciate “the complex, intricate, overlapping functioning of our sensorimotor systems, by deepening our understanding of how our nervous system binds us to our world, and showing how that world doubles back to shape us”(Robinson, 2015; 366).

Abstract /

Signs, in all their forms and manifestations, provide visual communication for wayfinding, commerce, and public dialogue and expression. Yet, how effectively a sign communicates and ultimately elicits a desired reaction begins with how well it attracts the visual attention of prospective viewers. This is especially the case for complex visual environments, both outside and inside of buildings. This paper presents the results of an exploratory research design to assess the use of eye-tracking (ET) technology to explore how placement and context affect the capture of visual attention. Specifically, this research explores the use of ET hardware and software in real-world contexts to analyze how visual attention is impacted by location and proximity to geometric edges, as well as elements of contrast, intensity against context, and facial features. Researchers also used data visualization and interpretation tools in augmented reality environments to anticipate human responses to alternative placement and design. Results show that ET methods, supported by the screen-based and wearable eye-tracking technologies, can provide results that are consistent with previous research of signage performance using static images in terms of cognitive load and legibility, and ET technologies offer an advanced dynamic tool for the design and placement of signage.

Keywords /

eye-tracking; signage; context

Given the significant contextual and cognitive issues related to the design and placement of signage, it is not surprising that research has found that not all signs, or even identical signs in different contexts, are equally successful in capturing the visual attention of potential viewers (Auffrey and Hildebrandt, 2017). How well a sign communicates its message reflects how design and placement accommodate and respond to the contexts of its surrounding built and natural environments. As such, tools and methods are needed that can inform design and regulation decisions by objectively assessing how well signs capture the attention of their intended viewers. This is especially true in complex visual environments, both outside and inside of buildings.

Capturing visual attention is important because it is a prerequisite for the conscious viewing required to achieve the response intended by sign owners. As such, how well a sign attracts visual attention is an important measure of its potential effectiveness in eliciting desired behavior. Yet, a shortcoming of signage design and regulation research is the limited use of objective measures of visual attention that includes the surrounding real-world visual context of signs (SRF, 2019; Garvey and Crawford, 2014). Further, signage research has been frequently directed at static, single-perspective images rather than the multi-viewer dynamic video analytics now available with wearable eye-tracking (ET) technology (Auffrey and Hildebrandt, 2017).

Consequently, the research presented here uses ET technologies to better understand the impact of signage design and regulation on the capture of visual attention across varying contexts. Of special interest is the use of ET methods and analytics from both screen-based and wearable ET technologies, which provide objective, multi-user performance measures in complex, real-world visual contexts.

RESEARCH DESIGN

To better understand how ET technology can be used to produce sign effectiveness metrics and therefore inform signage design and regulation, an exploratory research design was employed where both screen-based and wearable ET hardware and analytic software were used to collect and summarize data for two groups of research participants subjected to two discreet visual environments (see Figure 1). Exploratory research is useful in the development and refinement of research techniques, especially when new technology allows measurement and assessment that was not previously possible (Stebbins, 2001). In this regard, exploratory research methods can be especially useful to compare the new technology with earlier approaches and lay the groundwork that will establish the uses and limitations of new technology in specific applications.

As such, this research seeks to establish the uses and limitations of both screen-based and wearable ET



Figure 1 / Screen-based eye-tracking (L); Wearable eye-tracking (R).

hardware and software for signage research. ET studies are intended to capture conscious viewing and to solicit feedback from subjects about what they are looking at and thinking, so ET technology was applied to visually-complex real-world environments, with the intent to assess how it measures visual attention. These patterns would then be compared to established visual science factors. The specific exploratory approach used here uses descriptive, qualitative, and quantitative assessments of screen-based and wearable ET analytics for specified “areas of interest” (AOI) within photos and video capture. Heatmaps and gaze plots also were assessed. The research was conducted over two phases.

Phase 1: Screen-Based Eye-Tracking

The first phase used screen-based ET to document the specific visual elements which drew the visual attention of research participants. The specific exploratory approach uses descriptive, qualitative, and quantitative assessments of four ET analytics (mean time to first fixation, mean total fixation duration, mean fixation count, and mean total visit duration) for specific AOIs within four photos. Two of these images were from US college campus scenes that include buildings and walkways, a third was a street perspective of an urban art museum building in a US city, and the fourth was a street scene of commercial buildings in a Chinese city. A total of 45 college student volunteers participated; all were familiar with the US-based images and none were familiar with the Chinese scene.

Research participants were asked to observe a sequence of photographs while seated directly in front of a 24-inch high-resolution video display. The display was equipped with a Tobii Pro screen-based ET device with illuminators and a data processing unit. The processing unit collects data for individual participants and outputs aggregated metrics for image detection, 3D eye modeling and gaze mapping.

To begin, each subject went through the standard ET calibration and verification procedure, as specified by Tobii. The photographs were each displayed for three seconds, reflecting a predetermined “Time of Interest” (TOI). TOIs are defined interval of analysis used in ET research that allow for the organization of data according to time periods where meaningful behaviors

and events take place (Tobii, 2019). In this case, three seconds represent the period prior to conscious viewing, where subjects’ visual attention is directed to areas based on innate human response rather than conscious viewing.

In addition, AOIs were specified in each of the images, selected to assess how the measured eye movements were drawn to the visual elements considered triggers of the conscious viewing required for response. These include lines and edges, red/green and blue/yellow contrasts, intensity against context, and facial features within the visual environment. Based on the specific interests of the researchers, AOI boundaries were digitally selected; eye movement metrics would be calculated and exported for those elements. These metrics provide quantitative measures of how each AOI attracted visual attention across the participants. For this exploratory research, four metrics were measured for the selected AOIs:

1. Time to the first fixation, defined as the elapsed time between a start event until the first fixation occurs in each AOI,
2. Number of fixations, defined as the number of fixations that occur in each AOI,
3. Total fixation count for each AOI, and
4. Duration of fixation, defined as the elapsed time between the first gaze point and the last gaze point in the sequence of gaze points that makes up the fixation.

All four metrics are important indicators of how specific visual elements attract attention and ultimately trigger the conscious viewing required for response. As such, they can potentially inform decisions about signage design and regulation.

In addition to the fixation data for each of the AOIs, fixation heat maps were generated for the entire visual field and were used to identify those areas with the most focus of visual attention across multiple participants. The heat maps are a particularly useful tool for visualizing and understanding where fixations are most concentrated, comparing fixations across the different AOIs, and identifying those visual elements apart from the selected AOIs attracting visual attention.

Phase 2: Wearable Eye-Tracking

In the second phase, 22 student volunteers used Tobii Pro glasses, a wearable ET device, while walking in a defined space with observable signage. This device includes illuminators, a camera, and data collection and processing units for image detection, 3-D eye modeling, and gaze mapping algorithms. Like the screen-based ET, the images captured by the wearable ET are used to identify glints on the cornea and pupil. This information, together with a 3-D eye model, is then used to estimate the gaze vector and point for each participant (Tobii, 2019). In contrast to screen-based ET, wearable ET allows binocular coverage, with a full field of view and head tilt capabilities. Also, it avoids potential experimental bias resulting from a screen's size or pixel dimensions.

After a standard ET calibration and verification procedure, participants were instructed to walk in a defined space while wearing the glasses. In this case, the TOI was set at 60 seconds, recording a defined beginning and end of the visual occurrences over that period. Given the longer TOI period, the data collected reflects both pre-conscious (first three seconds) and conscious viewing.

Fixation and saccade data for the wearable ET device was exported for each participant. Fixations are the periods where the eyes are relatively still, holding the central foveal vision in place so that the visual system can take in detailed information about the focus of visual attention. Saccades are eye movements which move the fovea rapidly from one point to another. Based on the fixation filter and threshold, the fixation-saccade-fixation sequence can be computed in the Tobii analytical software. For example, if two gaze points are within a pre-defined minimum distance from each other or possess a speed below a defined threshold, they will be allocated to the same fixation (Tobii, 2019).

As with the screen-based ET, boundaries for AOIs were digitally selected so eye movement metrics would be calculated and exported for specific elements. For this exploratory research, three metrics were measured for the selected AOIs:

1. Time to the first fixation, defined as the elapsed time between a start event until the first fixation occurs in each AOI,
2. Number of fixations, defined as the number of fixations that occur in each AOI, and
3. Duration of fixation, defined as the elapsed time between the first gaze point and the last gaze point in the sequence of gaze points that makes up the fixation.

As discussed previously, these metrics are important indicators of how specific visual elements attract attention and trigger conscious viewing, thereby informing signage design and regulation decisions.

In addition to the AOIs, fixation data can be visualized as gaze plots and heat maps for the entire visual field. Gaze plots show the location, order, and time spent looking at specific locations in the visual field and heat maps are used to identify those areas with the most focus of visual attention across multiple participants. Both were generated identifying the visual pathways and fixations over the course of the TOIs for all 22 participants. A panoramic (360-degree) view is used for generating the gaze plots, heat maps, and AOIs.

FINDINGS

This research is part of an ongoing inquiry into the unique advantages of screen-based and wearable ET for studying signage within its context. The results reinforce and add to prior research about how the presence and placement of specific visual elements in signage and wayfinding design is important for the capture of both pre-cognitive visual attention and conscious viewing. As part of this exploratory research, the focus of this paper has been on four visual elements as part of signage design and placement. This includes proximity to edges and lines, red/green and blue/yellow contrast, intensity, and saturation of color, and use of facial features.

Proximity to Edges

The edge of an image can be defined as a boundary line reflecting changes in a pattern; edges define shape and necessarily attract visual attention. The results of this research show higher fixation metrics where linear

visual elements (edges or lines) appear near a background or nearby objects, such as with the wall edges or roof profiles against the sky. Fixation heat maps show higher levels of visual attention on signage along the edges of building façades. Further, aligning signage with edges potentially serves to facilitate wayfinding by directing a circulation flow based on visual cues.

Results from the screen-based ET analysis of an image with college campus buildings and walkway shows participants' visual attention was better captured by signage along the edge of a building façade rather than the larger, homogeneous façade system itself (see Figure 2). The results suggest that the alignment of signage with the façade edge may better serve wayfinding needs (in this case, building identification) and therefore facilitate pedestrian movement.

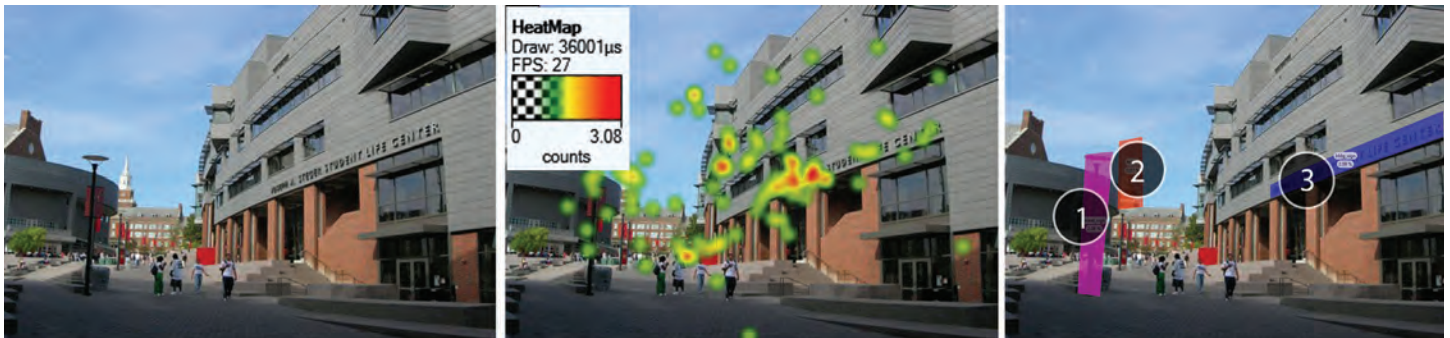


Figure 2 / College Campus Buildings and Walkway: Image (L), Heatmap (M), and AOIs (R).

For this image three AOIs were selected: the vertical banner hanging on the light pole, a building tower in the distance (a strong edge against the sky backdrop), and the block letter signage along the edge on the building façade. The ET fixation metrics for the AOIs show participants' attentions were most quickly focused on the block letter signage on the building (mean time to first fixation = 1.04 sec.) compared with the other AOIs (see Table 1). The block letter AOI also has the longest total duration (0.42 sec.), the largest fixation count (2.14), and the largest mean number of visits (0.46). These results suggest that signage placement near an edge increases its visual attention as a visual feature.

Table 1 / Area of Interest Eye-Tracking Metrics for Figure 2.

Areas of Interest	Time to First Fixation	Total Fixation Duration	Fixation Count	Total Visit Duration
1. Banner on light pole	1.5	0.29	2	0.31
2. Tower on distant building	1.33	0.38	1.33	0.4
3. Block letter signage on building façade	1.04	0.42	2.14	0.46

The wearable ET analysis provided additional results that could not be provided by the screen-based ET. The results suggest the importance of the placement of the signage relative to the expected viewing angle of potential viewers. The typical human vertical binocular coverage is approximately 50 degrees above the line of sight, and while wearable glasses allowed participants to tilt head freely, very few gaze points are recognized beyond this (see Figures 3 and 4).



Figure 3 / Art gallery panoramic image with “exit” sign indicated (above); Art gallery panoramic image with heat maps for eye fixations, N=22 (below).

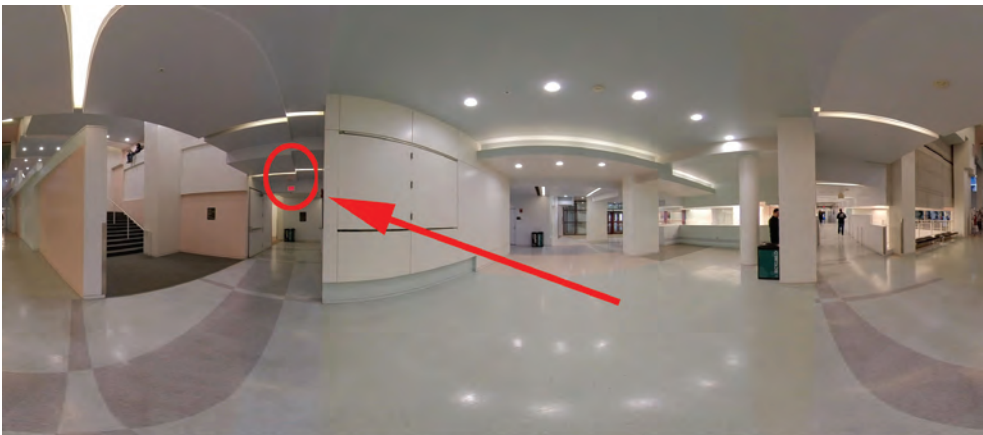


Figure 4 / Interior of campus building where “exit” sign is in a pattern-free wall, within the vertical binocular coverage, and reinforced by vanishing point.



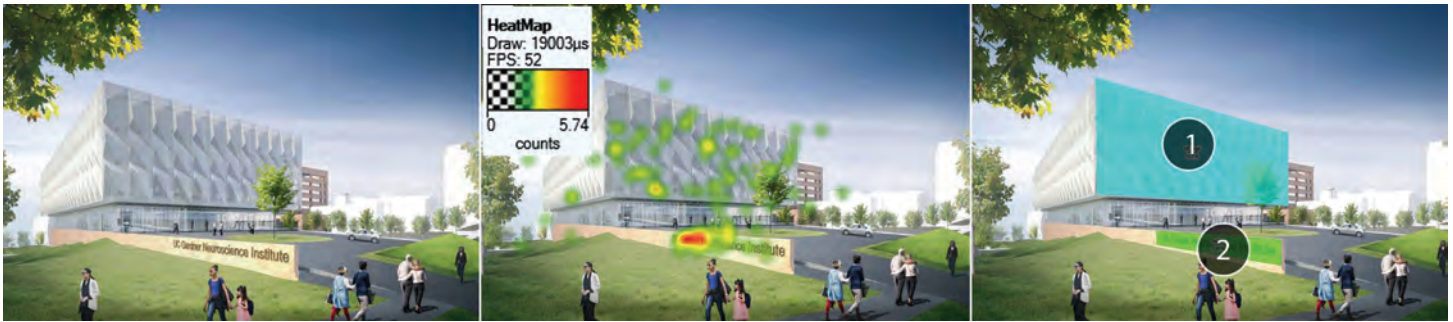


Figure 5 / Rendering of proposed college research center: Image (L), Heatmap (M), and AOIs (R).

Red/Green and Blue/Yellow Contrast

One of the images included in the screen-based ET analysis was used to explore the impact of red/green and blue/yellow contrast on visual attention. The image, a graphic rendering for a proposed college research building, included a low monument sign with color contrast and a large building façade composed of neutral gray material. Both the sign and façade were designated as AOIs for the analysis. Although the signage area represents only 1.5 percent of the image, compared to 11.7 percent for the building façade, the higher contrast of the sign elicited considerably more visual attention. The heat map shows that the number of fixations on the sign is higher than the building facade (see Figure 5).

The AOI metrics show higher values for the sign in all four categories: time to first fixation, total fixation duration, fixation count, and visit duration (see Table 2). Given that both AOIs contain substantial edge/line features, the higher contrast of the sign suggests an explanation for higher levels of visual attention on the signage.

Table 2 / Area of Interest Eye-Tracking Metrics for Figure 3.

Areas of Interest	Time to First Fixation	Total Fixation Duration	Fixation Count	Total Visit Duration
1. Building façade	0.72	0.93	4.17	1.04
2. Monument sign along wall	0.19	1.21	6	1.51

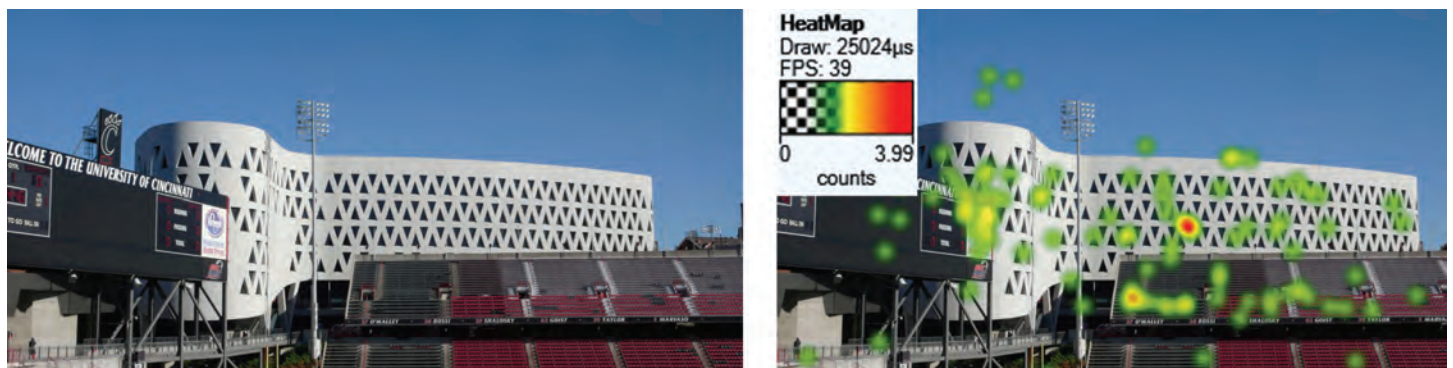


Figure 6 / Sign on scoreboard near high-contrast building façade: Image (L), Heatmap (R).

In comparison, an image with a low-contrast sign and a high-contrast building façade resulted in a different outcome (see Figure 6). The heat maps show how sharp triangular patterns and high black-white contrast on the building façade attracted more attention than the scoreboard mounted sign.

Intensity/saturation of color in context

The eye-tracking analysis was used to assess the impact of intensity and color saturation on visual attention. The pink lettering against a black background on the third level of an urban art museum building provided intensity and saturation of color that attracted higher levels of visual attention than the other elements of the building façade. The corner placement provides an edge to attract visual attention and allows visibility from two directions (see Figure 7). The heat map shows that the street sign and commercial lettering on the delivery truck received less visual attention.

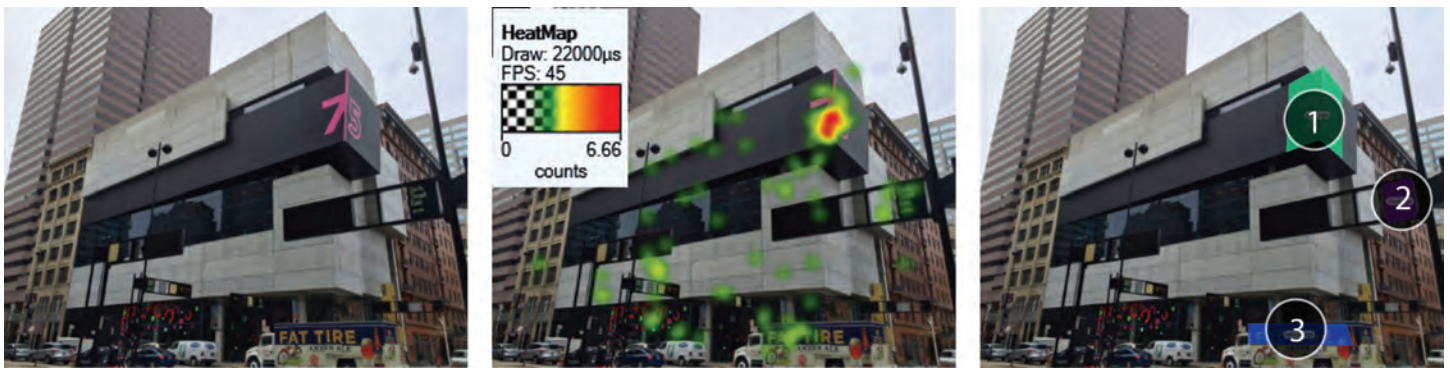


Figure 7 / Urban Art Museum Building: Image (L), Heatmap (M), and AOIs (R).

Table 3 / Area of Interest Eye-Tracking Metrics for Figure 6.

Areas of Interest	Time to First Fixation	Total Fixation Duration	Fixation Count	Total Visit Duration
1. Art museum building lettering	0.57	0.97	3.1	1.06
2. Street sign	1.65	0.65	1.5	0.66
3. Delivery vehicle lettering	2.39	0.46	2	0.5

Facial images in context

Human brains devote more areas to facial recognition than recognizing any other visual object (Sussman and Hollander, 2015). This reflects how our brains are hardwired to focus visual attention on other humans within sight. This priority is described by Chalup et al. (2010) as “face-a-tecture,” where the subconscious will perceive facial features as part of the structural design. This concept was explored using the screen-based ET to analyze an image with signage using a facial feature.

In a street scene of commercial buildings in a Chinese city, signs are installed on building façades and storefronts (see Figure 8). The scene contains strong

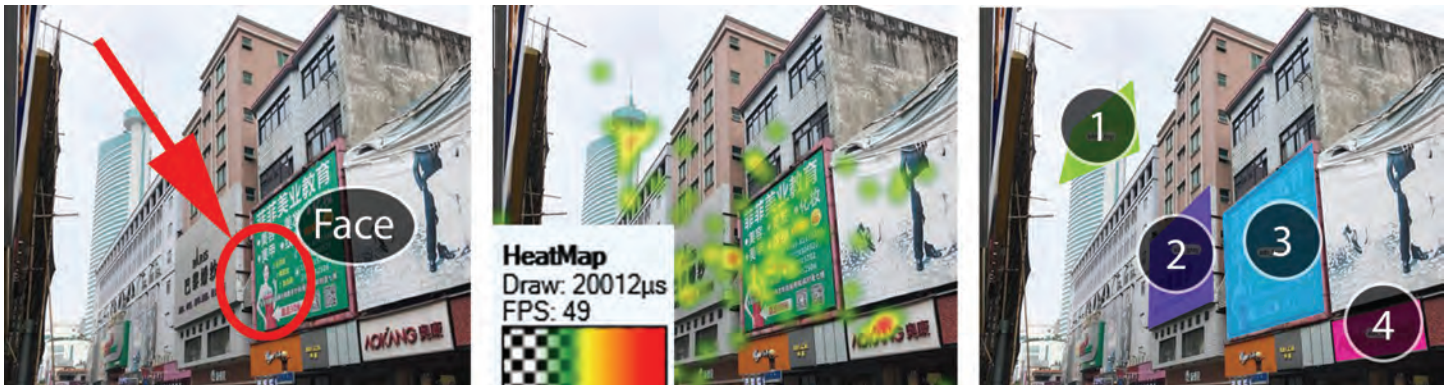


Figure 8 / Street scene of commercial buildings in China: Image (L), Heatmap (M), and AOIs (R).

Table 4 / Area of Interest Eye-Tracking Metrics for Figure 7.

Areas of Interest	Time to First Fixation	Total Fixation Duration	Fixation Count	Total Visit Duration
1. Distant building tower	1.43	0.34	1.4	0.38
2. Store name signage (far)	1.03	0.63	2	0.64
3. Signage with a human face	0.39	0.46	2.88	0.56
4. Store name signage (near)	1.34	0.43	1.75	0.45

edge features reflecting a vertical structure against the sky and a face image on signage (AOI #3). The screen-based ET metrics for the picture show that the signage with a human face received the first fixation and highest fixation count (see Table 4), however its total fixation duration is lower than other signs. Interestingly, the vertical tower also received visual attention, but with a longer “time to the first fixation,” suggesting the signage with facial features may have been a higher priority than the tower.



Figure 9 / Interior of building with a hanging poster containing human face, panoramic view: Image (above); heat maps (below).

The importance of facial features to attract visual attention also was explored using the wearable ET analysis. The resulting panoramic image and heat map of the interior of a college building show the relative visual attention on a hanging poster of a seated person whose face is clearly exposed (see Figure 9).

While the face on the poster is attracting eye fixations, the total number of fixations are fewer than the fixations on the seated persons below the poster. As with previous findings, it appears that visual attention is reduced if the facial features are outside the viewers' vertical binocular coverage. Although the wearable glasses allow people to tilt head freely, relatively fewer eye fixations are recorded outside the vertical binocular coverage.

CONCLUSIONS AND IMPLICATIONS FOR FUTURE SIGNAGE RESEARCH AND PRACTICE

The screen-based and wearable ET technologies are useful tools for assessing visual attention in complex

environments. While the implications for more efficiently promoting commerce and routine wayfinding are important, even more essential are the lessons for emergency egress from crowded buildings. The results of this exploratory work reveal both opportunities and limitations of ET research tools regarding visual communication challenges.

Clearly, ET technologies have demonstrated their usefulness to provide information that can be used to refine concepts throughout the design process and provide quantitative measures of the performance of signage design and placement given its context related to architectural forms, edges, contrasts, intensity, and facial patterns. The examples provided from this research have shown how quickly specific elements of signage and its surrounding visual context can attract viewers' attention and for how long that is sustained. The practical applications that emerge from this knowledge is useful, but more precise design clues are needed for complex, high-stakes environments. ET gaze plots, heat maps, and fixation metrics provide results that



Figure 10 / Interior of building, panoramic view: Gaze data from wearable ET (top); Gaze data from screen-based ET (bottom).

require and deserve further study to better uncover adaptive responses and optimize attention in unique environments. In short, the quantifiable ET analytics have the potential to improve visual attention by informing sign design as related to its particular context.

Given its potential and growing use as a research tool, the limitation of screen-based ET, that it does not consider the human binocular field of vision, must be recognized. Wearable ET data shows few head tilts and limited gaze points outside the vertical binocular coverage, even though ET research participants can tilt their heads freely. As a result, most of the gaze points have been concentrated at a middle height within the captured images, however, with screen-based ET there are more gaze points towards the upper part of the image because there is no consideration of head tilt (see Figure 10). The assumption is that the wearable ET results better reflect how real-world viewers survey a visual scene, though further research is needed to confirm this.

In conclusion, the five basic visual elements that attract visual attention in humans should be critical drivers for those concerned about visual communication. Further research using research tools from other fields, such as pupil dilation response, electroencephalography (EEG), and galvanic skin conductance, as adjuncts to ET analysis may provide additional insight and new perspectives for signage design. Such a combination of tools may better reveal how specific elements provoke emotional responses, whether calming or stressing. The benefits of these additional sensory data as interpretable metrics may open new opportunities for signage and wayfinding research that takes advantage of findings in neuroscience.

ACKNOWLEDGMENT

The research project is supported by the Strategic Collaborative/ Interdisciplinary Award of the University of Cincinnati. Thanks to the support from Professor Christopher Auffrey, students from ARCH7014 Fall 2019 semester, ARCH8001 Spring 2019 semester, and ARCH4001 Fall 2018 semester at the University of Cincinnati.

REFERENCES

- 3M Commercial Graphics Division. (2017). *3M Visual Attention Service Validation Study*. St. Paul: 3M Corporation.
- Auffrey, C. and Hildebrandt, H. (2017). Do motorists see business signs? Maybe. Maybe not. A study of the probability that motorists view on-premise signs. *Interdisciplinary Journal of Signage and Wayfinding*, 1(2).
- Chalup, S., Hong, K. and Ostwald, M. (2010). Simulating pareidolia of faces for architecture image analysis. *International Journal of Computer Information Systems and Industrial Management Applications*, 2, 262-278.
- Eberhard, J. (2009). Applying neuroscience to architecture. *Neuron*, 62(6), 753-56.
- Garvey, P and Crawford, R. (2014). *On-Premise Sign Research Review*. Alexandria: National Signage Research Foundation.

- Moore, R. (2011). *Improving Sign Effectiveness with Visual Attention Modeling Software*. St. Paul: 3M Corporation.
- Robinson, S. (2015). John Dewey and the dialogue between architecture and neuroscience. *Architectural Research Quarterly*, 19(4), 361–67.
- Sign Research Foundation (SRF). (2019). “Research Best Practices.” <https://www.signresearch.org/research-best-practices/>
- Stebbins, R. (2001). *Exploratory Research in the Social Sciences*. London: Sage.
- Sussman, A. and Hollander, J. (2015). *Cognitive Architecture: Designing for How We Respond to the Built Environment*. New York: Taylor and Francis.
- Sussman, J., Ward, J. and Hollander, J. (2018). “How biometrics can help designers build better places for people.” Common Edge, 05 April. <https://commonedge.org/how-biometrics-can-help-designers-build-better-places-for-people/>
- Tobii. (2019). “Working with Data.” tobiipro. <https://www.tobiipro.com/learn-and-support/learn/steps-in-an-eye-tracking-study/data/>