

Using eye-tracking to study designers' cognitive behaviour when designing with CAAD

Rongrong Yu¹, and John S. Gero²

¹ *School of Engineering and Built Environment, Griffith University, Southport, Australia*

r.yu@griffith.edu.au

² *University of North Carolina at Charlotte, USA and Krasnow Institute for Advanced Study, George*

Mason University, Fairfax, USA

john@johngero.com

Abstract: This paper presents a case study exploring the eye-movement of two architecture master students while using a CAAD tool. The students completed an architectural design task using CAAD software in a 60 minute design session. The “Think aloud” method was used to collect cognitive data while their eye movements were captured using eye-tracking equipment. The session was segmented and coded for visual location. Results of this exploratory study indicate differences when using a CAAD tool when designing compared to visual scanning of existing objects. These results, if generalizable, imply that CAAD software development should focus on space making as well as on object boundaries.

Keywords: Eye-tracking; CAAD; Protocol analysis; Design cognition

1. Introduction

Design is a purposeful, constrained, decision making, exploratory and learning activity (Gero, 1990). With computer-aided design tools increasingly used in the AEC industry, it is critical to explore how computer-aided design environments assist designers' activities. One of the many methods to explore designers' cognitive activity is using eye tracking technology. Eye movement data is quicker and more directly collected than much other cognitive data (Ware and Mikaelian, 1987). The user tends to look at where he or she would like to move before the actual movement (Jacob and Karn, 2003). Therefore eye-tracking technology is an effective tool for studying user's cognitive behavior. Previous research has focused on how designers look at static scene using eye-tracking technology (Gero *et al.*; Kaufman and Richard, 1969; Weber *et al.*, 2002; Yu and Gero, 2017). Marr (Marr, 1980; 2010) laid the cognitive foundations for visual processing using edge detection, image intensity changes, and object recognition of computational design.

However, there is a lack of empirical evidence regarding designers' activities during the dynamic design process in computational design environments. Do designers behave in the same way as people

who are looking at existing objects when they are designing? There is a gap in our knowledge about designers using CAAD tools. What can this knowledge tell us about CAAD tools?

To address this knowledge gap, this exploratory study focuses on how a computer-aided tool is used by two designers when they are doing architectural design tasks by studying their eye-movements. Two master of architecture students from Harbin Institute of Technology in China were recruited to participate in the experiment. The design task was to produce an architectural design in a computer-aided design environment in 60 minutes. Using both protocol analysis and eye-tracking, the students' design activities in a computer-aided design environment were captured. The remainder of the paper introduces eye-tracking and protocol analysis and then briefly describes the experiment, which is in the form of an exploratory case study.

2. Background

2.1. Eye-tracking

Early work of eye tracking technology can be traced back to Dodge and Cline (1901) who developed the first precise eye tracking technique, which used light reflected from the cornea. When it comes to the relationship between eye movement and perception of static pictures, one of the few studies on this topic related to architecture was conducted by Weber, Choi and Stark (2002). They collected eye tracking data as participants were asked to look at three-dimensional models or photographs of models, of an architectural space. 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 centre; 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. Arnheim (1974) also found that people tends to read pictures from left to right. This was confirmed in later results by Arnheim (1985). The Weber et al study also concluded that fixations did not vary significantly when viewing the physical model compared with a photograph of the model, with the exception of the foreground, which attracted greater attention in the physical model. Additionally, their results also suggest that there were significant differences between the fixations and saccades of architects and non-architects. Jacob and Karn (2003) studied eye movements while utilizing a user interface within human-computer dialogue, they pointed out that the number of fixations (Fixation Frequency) on each area of interest can be a measure of the importance of the information content of that area compared to other areas. Kaufman and Richard (1969) measured eye fixation times in pre-defined parts of figures, the results show that the centre of gravity is an attractor as are the edges and corners. These researches mostly discuss participants' eye movement when they look at static pictures. However, currently there is little research on exploring eye movement during a dynamic design process.

2.2. Protocol analysis

Protocol analysis is a method for turning qualitative verbal and gestural utterances into data (Ericsson and Simon, 1993; Kan and Gero, 2017). According to Akin (1986), a protocol is the record of the behavior of designers using sketches, notes, videos or audio. After collecting the protocol data, a coding scheme is applied to categorize the data, enabling detailed studies of the design process in the chosen design

environments. Protocol analysis has become the prevailing experimental technique for exploring the way in which the design process is understood cognitively (Cash et al., 2016).

As one of the main design ontologies, Gero's FBS ontology (Gero, 1990) has been applied in many cognitive studies (Gero and Tang, 1999; Kan and Gero, 2005; Song, 2014). Researchers argue that it is potentially capable of capturing most of the meaningful cognitive aspects of design when used as a system of coding the transitions between design issues into eight design processes formulation, analysis, evaluation, synthesis, and reformulation I, II and III (Gero, 1990). The FBS ontology, Figure 1 shows the three classes of ontological variables: Function (F), Behavior (B) and Structure (S). Function (F) represents design intentions or purposes; behavior (B) represents the object's derived behavior (Bs) or expected behavior from the structure (Be); and structure (S) represents the components that make up an artifact and their relationships. The ontology as the basis of a coding scheme includes two additional design issues that can be expressed in terms of FBS and therefore do not require an extension of the ontology. These are requirements (R) and descriptions (D). The first of these represents requirements from outside the design itself and the second, descriptions, refers to the documentation of the design, Figure 1. Gero's FBS coding scheme is structured at an ontological level and includes focusing on the intentions of the designer, and it has been claimed as a potentially universal coding scheme adaptable to different design environment (Kan and Gero, 2009). Dinar *et al.* (2012) propose a refinement of the FBS ontology by adding a hierarchical subdivision of the original FBS design issues, to better represent problem formulations of designers with different expertise levels.

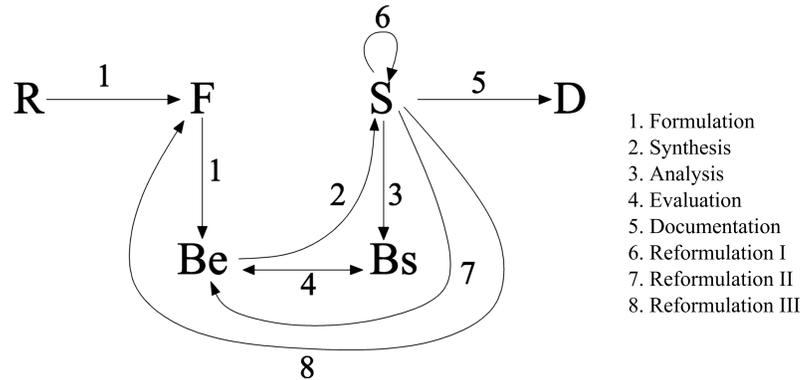


Figure. 1: The FBS ontology (after (Gero, 1990)).

3. Experiment – An exploratory case study

In this exploratory case study experiment, two architecture master students from Harbin Institute of Technology, China, were recruited as participants. Architecture master students were selected because they should have sufficient architectural design experience and knowledge to execute a complete design process. In the experiment, designers were required to complete a defined architectural design task using computational modelling software, which in this experiment is Sketchup. The participants each have more than 5 years' experience using Sketchup.

During the experiment, the designers' activities and their verbalization were video-recorded and the recorded data subsequently used for protocol analysis. The designers were given 60 minutes for the design session. The design task was a conceptual design for a high rise building with specific functions listed and located on a pre-modelled site, which was provided to the designers. During the experiment,

the designers were as not allowed to sketch manually so that almost all their actions occurred on the computer to ensure that the design environment was purely within the screen. This design environment is close to a real-world design environment at this stage of the design. The designer's eye-movements are recorded using eye-tracking equipment – in this study, we used Tobii Studio. Figure 2 shows the experiment set up.



Figure. 2: Experiment setting

5. Results

5.1 General results

This protocol study employed an integrated segmentation and coding method. The segmentation and encoding process are based on the “one segment one code” principle (Kan and Gero, 2017). This means there are no overlapped codes or multiple codes for any segment. If there are multiple codes for one segment, the segment will be further divided.

Two sets of coding schemes were applied in this research, the first one is based on the designer's verbalisation/activities, which is founded on the FBS ontology. The second coding scheme is based on the designer's eye-tracking data, which gives the location of her gaze. Table 1 shows the results of the segmentation.

Table 1: General results

	Total Segments (FBS)	Total Segments (Eye-tracking)	Design Time (Mins)
Designer 1	360	1191	49.03
Designer 2	357	1118	36.29

In order to obtain a more detailed articulation of what designers focus on when modelling geometry, we further divide Structure (S) into three sub-categories: Ss (site), Sb (base), and St (tower). The design issue distribution for the designer, derived from applying the FBS-based coding scheme is presented in Table 2. From the data in Table 2 we can see that both designers spent most of their design effort on structure (S) related activities. Among structure both designers expended more cognitive effort on St (53.1% and 23.0%), which is the geometry of tower, followed by Sb (base, 14.2% and 20.2%) and Ss (Site, 9.4% and 4.8%). After structure (S), the second highest effort was expended on Bs (13.1% and 24.7%), which is related to the examination of the design as it is generated.

Table 2: Design issue analysis

	R (%)	F (%)	Be (%)	Bs (%)	Ss (%)	Sb (%)	St (%)
Designer 1	1.9	3.6	4.7	13.1	9.4	14.2	53.1
Designer 2	1.1	7.0	19.1	24.7	4.8	20.2	23.0
Mean	1.5	5.3	11.9	18.9	7.1	17.2	38.0

The coding scheme for the eye tracking data was developed based on an examination of the eye gaze data. It used the location the eye gazes in the design, Table 3.

Table 3: Coding scheme for eye tracking data

	Code
Site planning	Surrounding of the site
	Corner
	Edge
	Face
Base	Corner
	Edge
	Face
Tower	Corner
	Edge
	Face
Menu	M

5.2. Cumulative analysis of protocol and eye-tracking data

The cumulative occurrence of design issues represented by the codes is determined from the segmentation and coding of the design session (Gero and Kannengiesser, 2012). The cumulative occurrence graph provides a basis for the qualitative evaluation of the expenditure of design effort, which correlates with cognitive effort, during a design session and is shown in Figures 3 and 4. From a visual inspection of the results in these figures we can see that for Designer 1 the predominant activity was associated with the tower (St). The designer worked on the base (Sb) at the beginning of the session and then no longer addressed the base for the remainder of the session. While for Designer 2, he worked on the base (Sb) first, and then focused on the tower (S-t). This shows the different sequence of designing. For both designers St and Bs have increased activity towards the end of the design session (indicated by an increase in slope). This means that the designer worked increasingly on the geometry of tower until the end of design session and kept examining the geometry. We can infer that the designer worked on the base and site of the building more at the beginning rather than in the end. For both designers F increases faster at the beginning of the design session. This suggests that the designer tended to consider the function related issues at the beginning of the design session. This agrees with results of previous cognitive research on design processes (Gero and Kannengiesser, 2012; Yu et al., 2013).

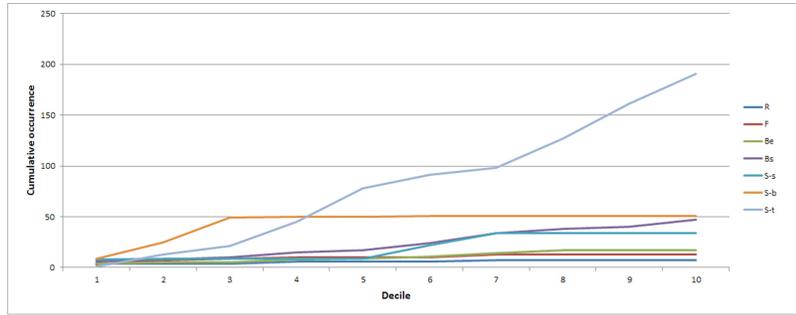


Figure.3: Cumulative analysis of FBS coding (Designer 1)

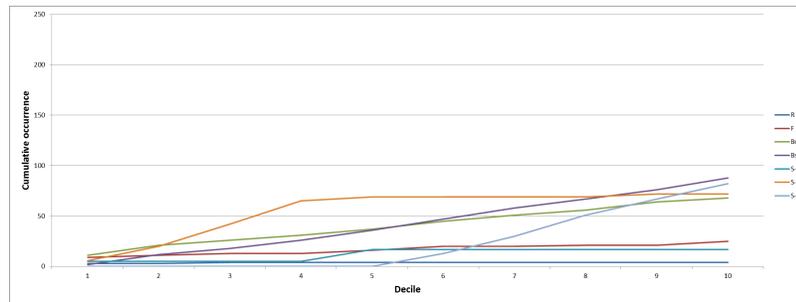


Figure.4: Cumulative analysis of FBS coding (Designer 2)

The cumulative analysis of eye tracking data is produced by segmenting and coding the eye gazes on façades, edges and corners of the building, Figures 5 and 6. From the results in these figures we can see that for both designers, their eye gazes focus more on the façades (a total of 723 gazes and 733), followed by edges (245 gazes and 241 gazes) and then corners (144 gazes and 55 gazes). In architectural design, shape and space are two fundamental primitives usually considered by designers (Moore and Allen, 1977; Ching, 2014). Previous research shows that designers’ eye-movement tends to focus more on the edge of static images rather than space (Weber et al., 2002). Results from the current study of dynamic design process suggest the opposite. This means when Designer 1 was designing, her attention is more on the spaces rather than edges, and rarely on corners.

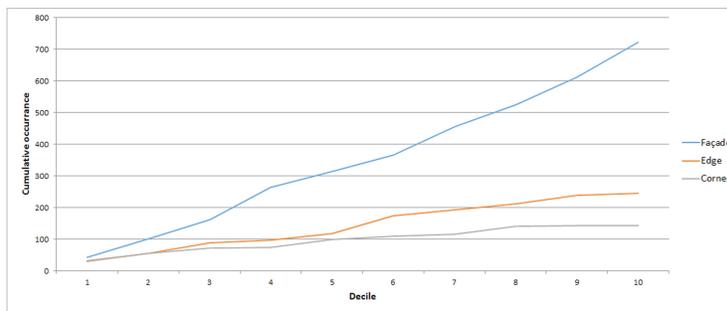


Figure. 5: Cumulative analysis of the eye-tracking gazes on façades, edges and corners (Designer 1)

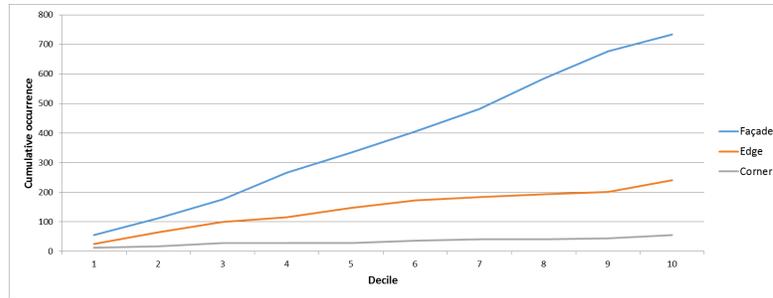


Figure. 6: Cumulative analysis of the eye-tracking gazes on façades, edges and corners (Designer 2)

Figures 7 and 8 show the cumulative analysis of eye tracking data of the two designers. From the results in these figures we can see for both designers that Tf gazes keep increasing towards the end of design session (up to 541 and 360). This means that designers worked on the façade of the Tower throughout the session, with an increase of attention towards the end of design session. For Designer 1, the attention of the Tower façade dominates the entire session and for Designer 2, Tf dominates the second half of the design session. This is followed by attention to the Base façade. The designers' gaze was very rarely focused on the screen menus. This is possibly because these designers used shortcuts.

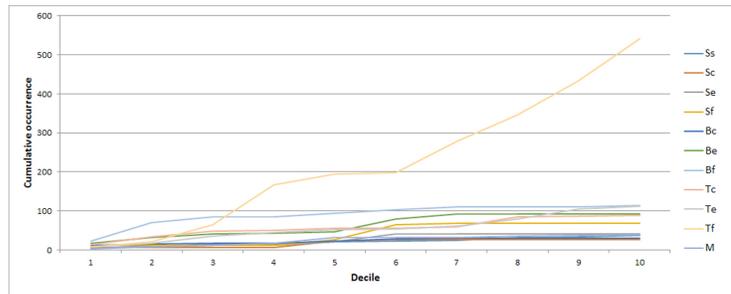


Figure.7. Cumulative analysis of eye-tracking coding (Designer 1)

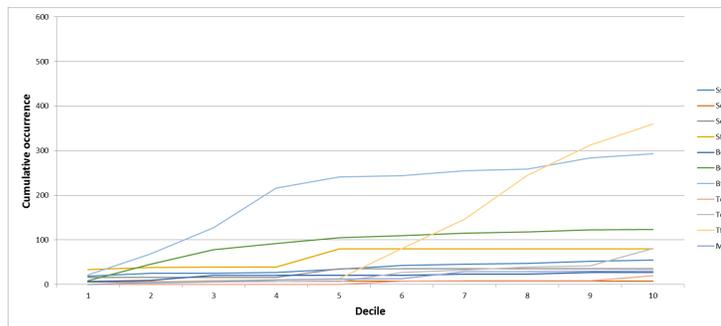


Figure.8. Cumulative analysis of eye-tracking coding (Designer 2)

5.3. Heatmaps of eye movement

From the heatmaps of the eye movement, Figures 9 and 10, we can see that most of the eye-tracking gazes are in the middle of the screen, some of the locations are towards the left-hand side. Only rarely do eye gazes appear on the right-hand side. This agrees with Arnheim (1974) and Weber et al. (2002) that participants' focus usually on the left hand-side of a picture. Another finding is that most of the eye tracking gazes are focused on the model, and only a few of them on the menu, this may be because the participant used shortcuts in Sketchup.

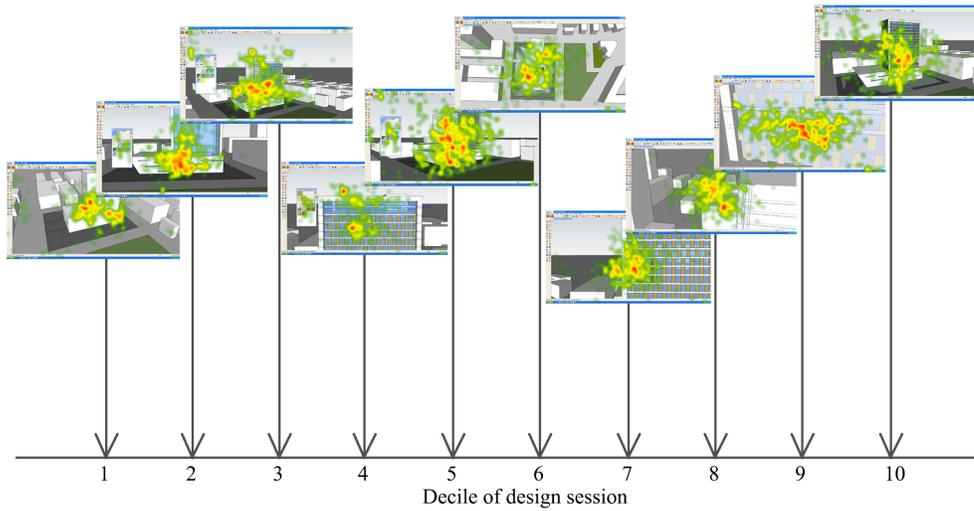


Figure 9: Heatmaps of the eye-tracking during design in deciles. Red-yellow-green represents the gradation of time focused in that region, with red the highest, yellow the middle and green the lowest (Designer 1).

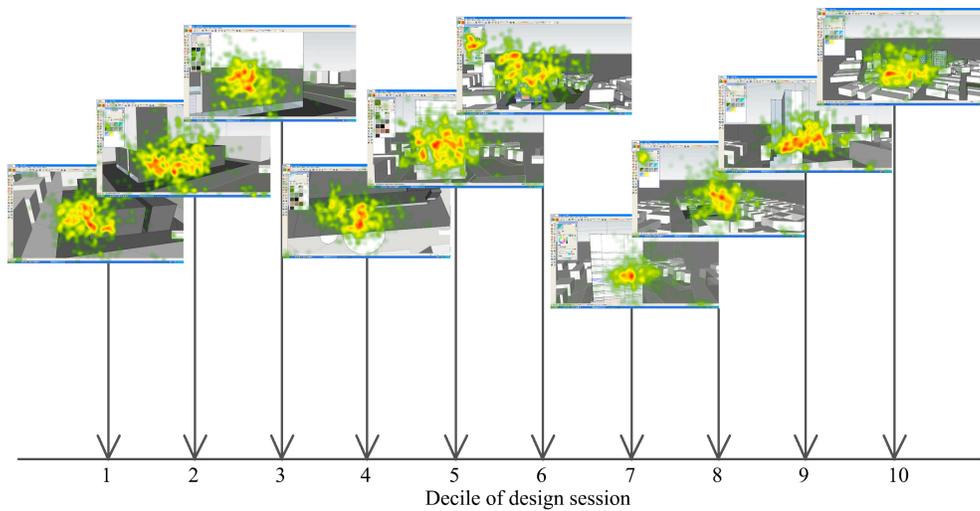


Figure.10. Heatmaps of the eye-tracking during design in deciles. Red-yellow-green represents the gradation of time focused in that region, with red the highest, yellow the middle and green the lowest (Designer 2).

6. Conclusion

This paper has presented the results from an exploratory case study on how a computer-aided modelling design tool is used by two designers during their design process using eye-tracking to capture their gazes. From the case study, we have following findings:

Firstly, the designers spent more design effort focusing on the façades and relatively less on the edges and corners. This result differs from previous studies on visual behavior that found edges are the centre of attraction rather than surfaces (Weber et al., 2002). This suggests a hypothesis to be tested: Do designers when using modelling tools focus on facades and spaces more than on edges and corners?

If this hypothesis is supported by more generalizable results it suggests that CAAD tools should have more functionality related to the making of shape and space rather than only on shape. The results in this case study which imply that designers focus more on the space rather than edges is the opposite of when they look at static pictures where they focus on edges (Marr, 2010). This suggests that designing, which is a generation process rather than an object recognition process and the results from object recognition are not directly applicable to designers (Ullman, 2000; Cyganek 2013).

Secondly, the designers examined the model constantly towards the end of design session. The checking of the existing model is essential during design process (Yu et al., 2013). This suggests a hypothesis to be tested: Is it beneficial for CAAD tool design to focus on developing the navigation tools which can help designers examine their models more easily?

Thirdly, the case study shows that these experienced designers' eye gazes focused primarily on the model with only a few eye gazes on the menu of the interface. This suggests a hypothesis to be tested: Is it beneficial to the designer that for experienced users of the software the menu design can be simplified and take less space.

Finally, most of the designer's eye gazes focused on the middle of the screen or towards the left. Marr (2010) suggests that 2D dimensional sketch is more viewer centred in terms of geometry, while 3D dimension images are more object centred, which includes volumetric primitives. This provides the basis of a final hypothesis to be tested: Is it beneficial to the designer if the interface leaves more modelling space in the middle and left, and locates menus on the right?

Results of this study provide the foundation for hypotheses that can be tested in more generalizable experiments. The benefit of applying eye-tracking technology in design cognition research is that it can provide biometric evidence about where designers look, and link this information to what they think. Prior to this research, designers' cognitive behaviour during their design process had not often been related to their eye-movement. The method proposed from this study opens broader possibilities of cognitive research from a biometric perspective. The results of this research are necessarily limited by the very small number of participants. However, due to the nature of protocol analysis (time consuming however can achieve detail analysing of designers' cognitive thinking) and eye-tracking analysis (large amount of data generated for eye-gazing analysis), the depth of the analysis provides robust evidence of designers' cognitive behaviour coupled with their biometric response during the design process. The results of a broader range of experiments could form the basis for guidelines to CAAD software designers and be beneficial for both designers and software developers.

References

Akin, O. (1986) *Psychology of Architectural Design* ed., Pion, London.

- Arnheim, R. (1974) *Art and visual perception: A psychology of the creative eye*, ed., University of California Press, Berkeley, Los Angeles, London.
- Arnheim, R. (1985) *The power of the center*, ed., University of California Press, Berkeley.
- Cash, P., Stankovic, T. and Storga, M. (eds.) (2016) *Experimental Design Research*, Springer.
- Ching, F. (2014) *Architecture: Form, Space, and Order*, 4th ed., Wiley.
- Cyganek, B. (2013) *Object Detection and Recognition in Digital Images: Theory and Practice*, ed., Wiley.
- Dinar, M., Maclellan, C., Danieleescu, A., Shah, J. and Langley, P. (2012) Beyond Function-Behavior-Structure, *Proceedings of Design Computing and Cognition (DCC) '12 Conference*, Texas, USA, 511-527.
- Dodge, R. and Cline, T. S. (1901) The angle velocity of eye movements, *Psychological Review*, 8, 145-157.
- Ericsson, K. A. and Simon, H. A. (1993) *Protocol analysis: verbal reports as data*, ed., MIT Press, Mass.
- Gero, J. and Tang, H.-H. (1999) *Concurrent and Retrospective Protocols and Computer-Aided Architectural Design*, in Z. Xie and J. Qian (eds.), CAADRIA1999, Shanghai, 403-410.
- Gero, J. S. (1990) Design prototypes: a knowledge representation schema for design, *AI Magazine*, 11(4), 26-36.
- Gero, J. S. and Kannengiesser, U. (2012) *Commonalities across Designing: Empirical Results*, in J. Gero (ed.), *Proceeding of International conference of Design Computing and Cognition (DCC'12)*, Berlin: Springer, College Station, 285-302.
- Gero, J. S., Shields and Yu, R. How veridical are different modalities of digital representation? , *Living Systems and Micro-Utopias: Towards Continuous Designing, Proceedings of the 21st International Conference of the Association for Computer-Aided Architectural Design Research in Asia, CAADRIA 2016*, edited by S. Chien, S. Choo, M. A. Schnabel, W. Nakapan, M. J. Kim and S. Roudavski, 861-870.
- Jacob, R. and Karn, K. (2003) Eye tracking in Human-Computer Interaction and usability research: ready to deliver the promises, in J. Hyönä, R. Radach and H. Deubel (eds.), *The mind's eye: cognitive and applied aspects of eye movement research*, 573-605.
- Kan, J. W. T. and Gero, J. S. (2005) *Can Entropy Indicate the Richness of Idea Generation in Team Designing?*, in A. Bhatt (ed.), CAADRIA 2005, New Delhi, India, 451-457.
- Kan, J. W. T. and Gero, J. S. (2009) Using the FBS ontology to capture semantic design information in design protocol studies, in J. McDonnell and P. Lloyd (eds.), *About: Designing. Analysing Design Meetings*, Taylor & Francis, New York, 213-229.
- Kan, J. W. T. and Gero, J. S. (2017) *Quantitative Methods for Studying Design Protocols*, ed., Springer.
- Kaufman, L. and Richard, W. (1969) Spontaneous fixation tendencies for visual forms, *Perception and Psychophysics*, 5, 85-88.
- Marr, D. (1980) *Theory of edge detection*, *Proceedings of the Royal Society of London B: Biological Sciences*, 187-217.
- Marr, D. (2010) *Vision*, ed., MIT Press, Cambridge.
- Moore, C. and Allen, G. (1977) *Dimensions: Space, Shape & Scale in Architecture*, ed., McGraw-Hill.
- Song, T. (2014) *Expert Vs. Novice: Problem Decomposition/Recomposition in Engineering Design*, Utah State University, Logan, Utah.
- Ullman, S. (2000) *High-Level Vision: Object Recognition and Visual Cognition*, ed., MIT Press.
- Ware, C. and Mikaelian, H. T. (1987) *An evaluation of an eye input*, ACM CHI+GI'87 Human Factors Conference, ACM Press, New York, 183-188.
- Weber, R., Choi, Y. and 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.
- Yu, R. and Gero, J. S. (2017) Exploring designers' cognitive load when viewing digital representations of spaces: A pilot study, *Research into Design for Communities, Volume 1 (Proceedings of ICoRD 2017)*, 457-467.
- Yu, R., Gero, J. S. and Gu, N. (2013) *Impact of using rule algorithms on designers' behavior in a parametric design environment: Preliminary results from a pilot study*, *Proceedings of the 15 th International Conference on Computer Aided Architectural Design Futures (CAAD FUTURES 2013)*, Shanghai, China, 13-22.