

# Processing static and dynamic diagrams: Insights from eye tracking

Richard Lowe<sup>1</sup> [r.k.lowe@curtin.edu.au](mailto:r.k.lowe@curtin.edu.au);

Jean-Michel Boucheix<sup>2</sup> [Jean-Michel.Boucheix@u-bourgogne.fr](mailto:Jean-Michel.Boucheix@u-bourgogne.fr)

<sup>1</sup>Curtin University, Australia; <sup>2</sup>University of Burgundy, France

Empirical studies of static and dynamic diagrams have traditionally collected outcome data indicating the effectiveness of these depictions with respect to comprehension and learning. Because outcome measures alone provide limited insights as to why diagrams are or are not effective, there has been growing interest in studying how people process these depictions. In some cases, the aim of this research is to develop principled approaches for guiding diagram design while in others it is to devise strategies that could support users. This paper presents a selection of examples from varied content domains illustrating how eye tracking data can be combined with other measures to probe how users interact with diagrams. Descriptions of the systems used in these combinations will be described and the synergies between eye tracking and the other measures explained. The illustrations are selected from studies in which the goals ranged from exploring the effects of cueing to comparing visual and haptic search. These different examples show that approaches used for analysing and interpreting eye tracking data need to be carefully matched to the specific goals of individual studies. We conclude with recommendations for using eye tracking as an adjunct to other approaches for gathering diagram processing data.


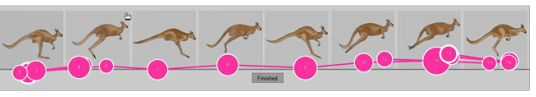
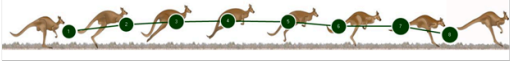



## Introduction

In recent years, advances in non-intrusive technologies for eye tracking have led to far wider adoption of this approach for scientific research. The application of eye tracking to research on diagrams has been prompted by an increasing interest in the way human users process these visualizations during comprehension and learning. Although eye tracking cannot provide direct evidence about cognitive activity, its use in combination with other measures can indicate mechanisms that may be involved in diagram processing. In this paper, we summarise a number of studies that have used eye tracking in this adjunct role. Our goal is to demonstrate a range of ways in which eye tracking can be allied to complementary measures in order to develop a more rounded understanding of how diagram processing occurs. The examples given show the importance of matching the approach taken in analysing eye tracking data to the specific aims of particular investigations.

## Effectiveness of Presentation Formats

Lowe, Schnotz and Rasch (2010) compared the effectiveness of different ways of preparing learners for the task of correctly arranging eight randomly ordered static pictures of a kangaroo hopping cycle. In the simultaneous condition, learners were prepared by showing them all eight pictures together in a correctly sequenced row for a period of eight seconds. In the successive condition, these pictures were presented one after another, with each being exposed for one second. In the animated condition, the eight picture sequence was presented repeatedly at 12 frames per second for a total of eight seconds. The arrangements produced

during the sequencing task were scored according to how closely they corresponded with the correct sequence. Eye tracking was used with a sample of participants to obtain data about how the three types of presentation were processed. It was hypothesised that the three groups would differ in where their visual attention would be directed. In the simultaneous and animated conditions, it was expected that participants would attend to the separation between the kangaroo's feet and the ground in order to use the kangaroo's elevation as an indicator of the correct sequencing. However, in the successive condition they were expected to pay more attention to changes in the configuration of the kangaroo's body, an aspect that is emphasised when one static picture is replaced by another static picture.

	Learning	Output
Dynamic		
Simultaneous		
Successive		

**Figure 1.** Participants in successive group directed attention to region that was most informative with respect to changes in configuration of kangaroo's body

Eye tracking data using Areas of Interest (AOIs) that divided the display region into lower and upper sections confirmed that those in the successive group directed more of their attention to the part of the display where the most distinctive changes in body configuration occurred. The superior sequencing performance of those in the successive condition was attributed to the distinctiveness of the inter-picture body configurations and the relations between one picture and the next. The types of sequencing errors produced in the simultaneous and animated conditions were consistent with the areas where participants in these groups tended to direct their attention. For example, compared with the successive condition, there were more cases where a picture that should have been in the upward part of the hopping cycle were misplaced to the corresponding position of the downward part of the cycle. This substitution suggests that these participants used only the elevation of the kangaroo's feet but not information about the distinctive configurations in the first and second halves of the hopping cycle that would have prevented confusion between the upward and downward sections.

### Between-picture and Within-picture Processing

Boucheix, Lowe, Groff, Paire-Ficout, Argon, Saby, and Alauzet (2011) used eye tracking to study the comprehension of diagrammatic public information messages. These diagrams were a purely visual format intended to provide information about traffic disruptions in French railway stations for people who cannot hear or understand normal loudspeaker announcements. For each of the five disruption messages tested, a series of four pictures were presented on a computer screen. These depictions followed the common script of events delivered in typical railway announcements as follows: the cause (*because of the bad weather*) the main events (*the train TGV number 67508 will be 15 minutes delayed*) and the

possible actions (*go to information desk or have a sit in the resting room*). The goals of the study were (i) to study the potential of static and dynamic visual displays to quickly and effectively trigger a task-appropriate schema of the relevant events, and (ii) to examine which type of visual format would efficiently map to travellers' internal schemas for train disruptions events. Within each of the four main pictures constituting the message as a whole were internal picture components (such as a train) which could be either static or animated. Four presentation formats were compared: Animated sequential, Animated simultaneous, Static sequential, and Static simultaneous. Comprehension was assessed verbally by comprehension questions asked after each message.

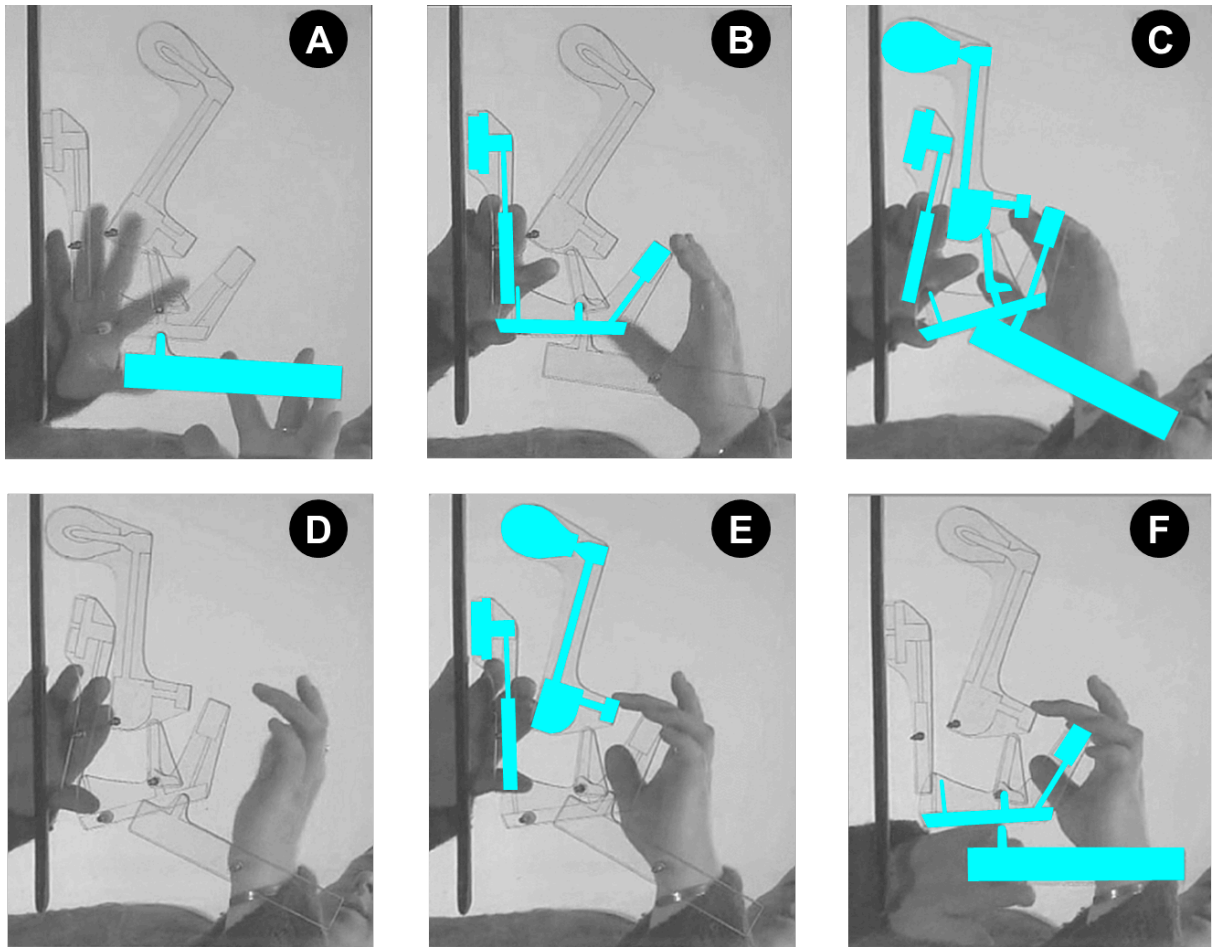


**Figure 2.** Areas of Interest capturing *between*-picture and *within*-picture aspects of processing revealed that sequential formats were tracked more regularly.

Eye movements were recorded for two reasons: (i) to analyze the extent to which participants did or did not follow the animation overall and (ii) to distinguish between-picture and within-picture aspects of processing. Results of the eye movement analyses suggested that superior comprehension for the animated sequential version was due to regular tracking behavior both between the four pictures and across components within the pictures. In the sequential formats, fixation durations of each picture decreased much more strongly and regularly from the first to the fourth pictures than in the simultaneous formats. This effect applied between pictures as well as within pictures, suggesting anticipation behaviors.

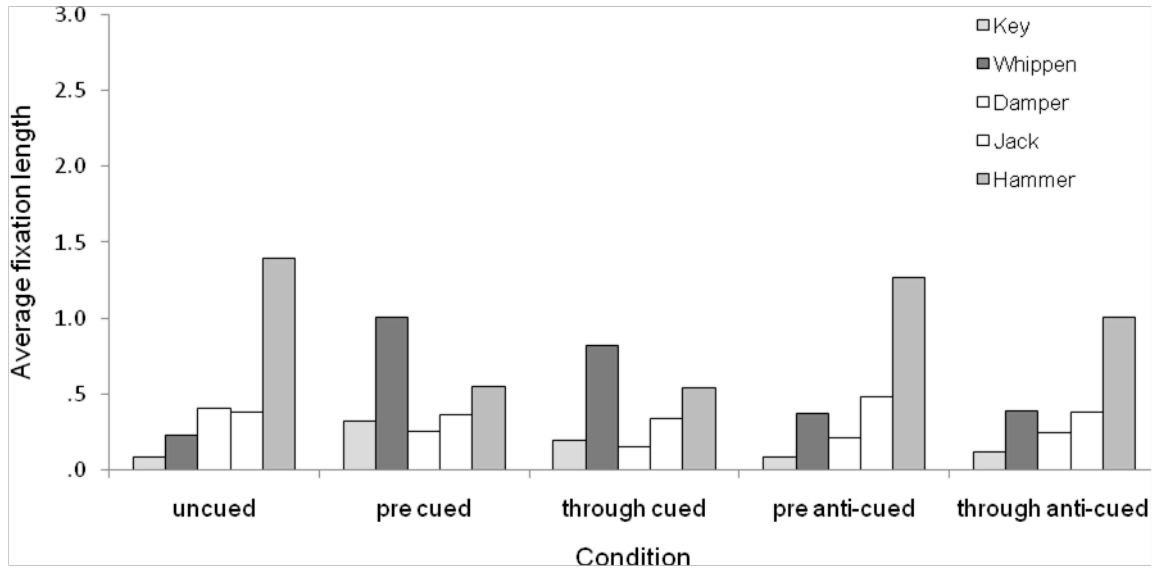
### **Animated Diagrams and Traditional Cueing**

Different methods of cueing components in a piano mechanism animation were compared in an approach where learners alternately viewed the animation then demonstrated what they had observed on a replica piano model (Lowe & Boucheix, 2011). Cues are supposed to improve processing of graphic displays by directing learner attention to high relevance aspects so that they are more likely to be noticed then internalized as part of the learner's developing mental model. However, while cues may be effective in static diagrams, results from experiments that use the same type of cueing in animated diagrams have been inconclusive.



**Figure 3.** Progressive demonstrations of piano mechanism operation following successive viewings of animation used to explore processes of constructing mental models

Eye tracking was used to gather evidence about why such cues may not function effectively in an animated context. This was done by constructing AOIs based on the various piano components and the paths they swept out during the mechanism's operation. The viewing/demonstration alternations were repeated ten times to explore how mental models of complex content progress with successive exposures. None of the cueing methods proved superior to no-cue controls.

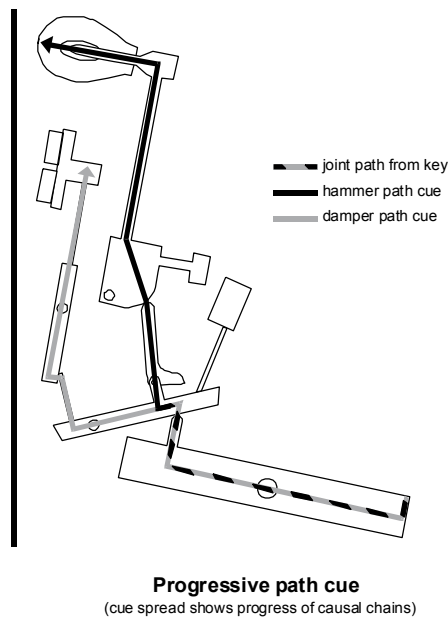


**Figure 4.** Relative duration of fixations on different components of the piano mechanism varied with cueing conditions (data for first animation exposure only)

The eye tracking data indicated that even in the few cases where cueing helped to capture attention on the first animation exposure, its influence all but disappeared with repeated viewings. On subsequent exposures, it appeared that the animation's dynamic characteristics were more influential in determining where learners directed their attention. For example, the hammer, which has the highest perceptual salience of all the piano mechanism's components, reliably attracted most attention, irrespective of the cueing used. This component has the most pronounced movement in the mechanism and so its dynamics have a far more powerful influence on attention direction than the static colour cues. It was also possible to infer differences in the intrinsic perceptibility of different piano components by comparing the eye tracking data for each area of interest. Another indication from the eye tracking data was that the pattern of exploration, as indicated by the fixations made in each AOI, changed across the series of viewings and demonstrations.

### Non-traditional Cueing

Most studies using eye tracking to investigate effects of cueing on the comprehension of animated diagrams have employed the global measure of total fixation time across the whole learning phase rather than more targeted measures. In a recent study, Boucheix, Lowe, Putri and Groff, (in press) used time-locked analysis of eye movement data to examine how promptly and faithfully learners "obeyed" cues presented during a task involving comprehension of a piano mechanism animation. The effectiveness of animations containing two novel forms of cueing that targeted relations between event units rather than individual entities was compared with that of animations containing conventional entity-based cueing or no cues. These relational event unit cues were respectively *progressive path* and *local coordinated* cues that signaled not only entities but also events along the causal chains spreading dynamically through the mechanism.



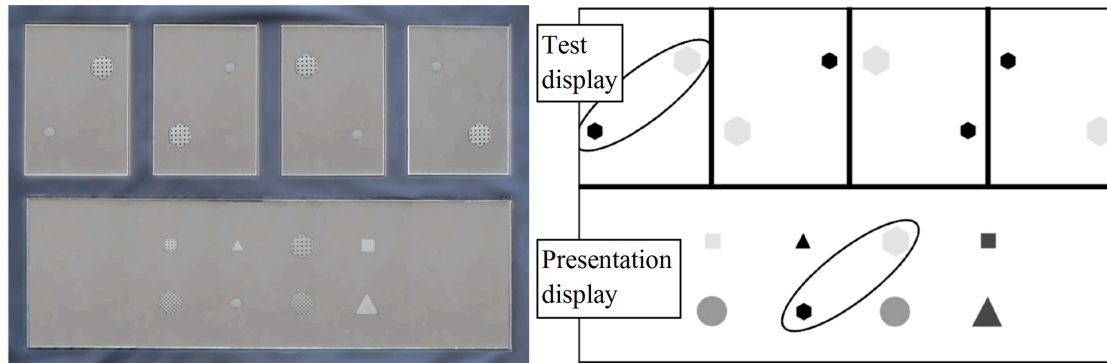
**Figure 5.** Eye tracking with time-locked analysis used to distinguish between engagement with and loyalty to event-based cueing approaches.

We distinguished two forms of cue obedience: (i) *engagement* – the cue’s initial capture of attention when it first appears, and (ii) *loyalty* – the further direction of attention to the cue beyond this initial capture. Cue engagement was operationalized as the time to first fixation, or as the number of fixations before a first fixation is made in the target area once the cued appears in that area (i.e., cue *entry*). Cue *loyalty* was operationalized as the relative amount of time spent viewing cued locations from the moment the cue appears in a specific AOI (*entry*) until it disappears of this AOI (*exit*). Results showed learners in the relational event cueing conditions fixated the target areas sooner than those in the entity cued group. Relational event unit cues not only directed learner attention more promptly to the target areas, they also resulted in a greater level of attention to those areas overall, improving cue loyalty. Comparison of the eye tracking data analysed at different scales indicated that cue obedience was partial rather than total. During the spread of the cues, an appreciable amount of time was also spent fixating information in uncued AOIs beyond the cued area. Despite cue obedience being only partial, relational event cues still produced superior learning. Rather than being a problem, this partial obedience is likely to be advantageous because it allows for the flexibility needed to build essential relations between material in cued and excluded regions.

### Visual versus Haptic Search

An investigation by Lowe and Keehner (2010) used eye tracking as a tool for comparing the processes people use when searching abstract visual and tactile displays for information. This study was conducted to help determine the potential of using haptic guidance for supporting more effective visual exploration of complex diagrammatic displays. The materials used were spatially equivalent displays composed of geometric shapes arranged in two rows of four shapes each. These displays were produced using either shading (for visual stimuli) or texture (for haptic stimuli) to provide distinctive surface renderings of items. Participants were asked to examine the display to determine which of four possible test items represented the correct configuration of subsets of shapes present in the display. In the visual condition, their search processes were characterised via eye tracking, while in the tactile condition, video recordings

made from a camera mounted below the transparent display were used. A key challenge in this study was to devise an approach to analysis that would allow the search processes to be compared across these very different types of data sets. This was done by deeming visual fixations to be foveations that corresponded with the functional visual field and haptic ‘fixations’ to be direct finger contacts with a display entity. Using this method, it was possible to quantify how speed and accuracy of visual search compared with these dimensions of haptic search.

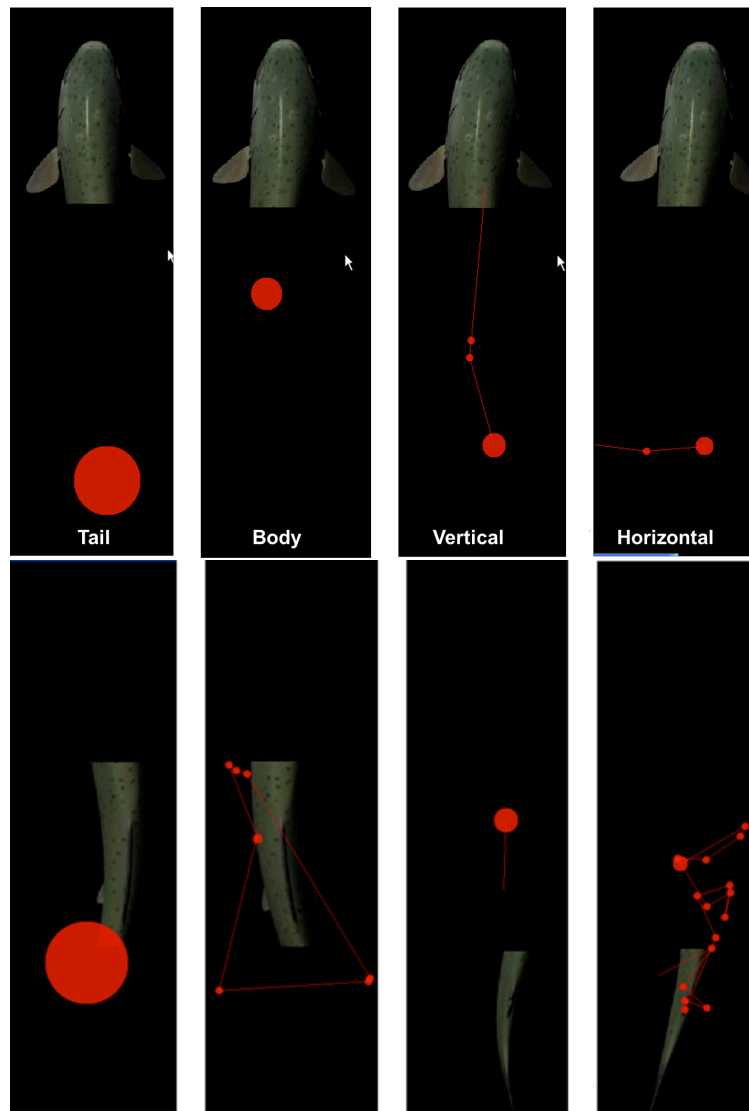


**Figure 6.** Tactile and visual versions of test display used to compare search strategies, efficiencies and effectiveness. Functional visual field used to define AOIs

Although visual search was many times faster than haptic search, the accuracy of these two search types was far more comparable. Detailed analysis of the visual exploration captured by the eye tracker and haptic exploration captured by the video suggest that the speed differences were due to both differences in the resolution of the two sensory systems and the number of fixations made per search task. The complementary use of eye tracking and video coupled with approaches to data analysis that were carefully tailored to suit the specifics of the investigation allowed meaningful comparisons to be made between two very different types of data sets.

### Inferring Missing Information

Researchers typically use eye tracking to study how learners look at different locations within a presented display. Boucheix and Lowe (2009) used eye tracking to examine the processing involved when learners were required to imagine information that was *not* presented on the screen. In this investigation, participants studied fish animations to learn the dynamic patterns involved in their locomotion. Initially, all participants were briefly exposed to a locomotion animation depicting the whole fish. Next, three out of four groups randomly selected from those participants studied a locomotion animation depicting only one section of the fish’s body: group 1, the head only; group 2, the body only; group 3, the tail only. Participants in these groups were instructed to use the partial information they had been supplied to imagine how the missing part of the fish moved. Those in the fourth group were controls provided with a whole fish locomotion animation. Time on task was the same for all groups. In the post test, participants performed a recognition task in which they were presented with animations that showed locomotion patterns that were either identical to that previously studied or differed from the original in varying degrees.



**Figure 7.** Eye tracking results suggesting participants' mental simulation of missing aspects of fish movements. Note fixations for both body parts and movements

Eye movements were recorded during the learning task as well as during the posttest. The collected data were analyzed according to areas of interest that corresponded to present or absent regions of the fish body. Eye tracking data collected during the learning time suggested that participants in the three partial fish animation groups were engaged in active mental simulation of missing aspects of the fish movements, looking intensely at the part of the screen where the relevant missing body part should have been located. For example, participants in the tail only group looked longer to the head location than participants of the control group who were presented the whole fish movements. Participants in the partial fish groups also produced better recognition performances than those in the whole fish movement group. This study shows that eye tracking is not limited to investigating how people process information that is explicitly provided in visualizations. It can also be a useful empirical tool for studying inferential processing where a depiction does not provide a comprehensive representation of the referent subject matter.



## Conclusion

The six studies summarized here demonstrate the variety of ways in which eye tracking can be used to help researchers understand what happens during diagram processing. This technique provides data that reflect the underlying reasons for variations in overall performance. In the first two studies, eye tracking was used to explore how varying the format in which a set of information was presented influenced how the viewer's attention was allocated in the display. In both cases, it suggested that the way information was distributed over space and time had important effects on which aspects of the presented content were extracted by the viewer. For example, eye tracking results suggested that the dynamic attributes of a presentation helped to determine whether the display was processed in a global manner or more analytically. Such insights allow us to judge which form of display is likely to be best suited to a particular task (sequencing information or comprehending instructions).

In studies three and four, the focus was not on the presentation format *per se*, but rather on the effect that including ancillary support has on performance. These two studies examined the effect of adding different types of cues to an animated diagram and used eye tracking to obtain finer grained information about processing. Rather than using broad brush measures such as the total number or duration of fixations, the eye tracking approach was applied in ways that gave snapshots of how visual processing changed over time. In the first case, eye tracking was repeated across successive inspections of the presented animation whereas in the second case, the eye tracking data from a single inspection sequence was subdivided into successive temporal chunks. This partitioning of the data provided additional insights into the way the viewer's processing changed and evolved over time as a mental representation of the depicted referent developed. It also helped to confirm the dominance of dynamic contrast over visuospatial contrast as an influence on viewers' direction of attention and to reveal how attention direction became more distributed as relational structures were built within the mental representation.

The final two studies depart even further from standard applications of eye tracking. In study five, the intention was not simply to determine which aspects of the display received most attention, but rather to investigate the patterns of inspection used to perform various tasks. The nature of these patterns helped to explain similarities and differences in visual and haptic search. Study six provided insights into how people process aspects of the referent content that are *not* actually depicted in a visual presentation. Extension of the fixation patterns into empty regions of the display suggested that viewers may use processes such as extrapolation to infer and predict the dynamic changes taking place in missing parts of the subject matter. This capacity for mental gap-filling has important implications for the comprehension of diagrams, an abstract form of depiction that is characterized by its high level of selectivity.

Our approach has been to use eye tracking in combination with other measures in order to benefit from the complementary perspectives the different data sets provide. Although eye tracking equipment comes with standard analytical software, it is sometimes better to work with raw data rather than rely only on the routine types of analysis that are provided. Eye tracking data are of limited value to diagram researchers on their own. They need to be complemented not only by other data collection approaches, but also well-developed theoretical modeling of how human diagram processing is considered to proceed. One of the most important considerations in applying eye tracking technology to diagrams research is that the specific approaches used are well matched to the purposes of the investigation.

## References

- Boucheix, J.-M., & Lowe, R.K. (August, 2009) *When less is more: Eye tracking comparison of dynamic inference during viewing and mental simulation*. Paper presented at the 13<sup>th</sup> European Association for Learning and Instruction Conference Amsterdam, Netherlands.
- Boucheix, J.-M., Lowe, R. K., Putri, D., & Groff, J. (in press). Learning from complex animation: Relational event unit cueing fosters efficient interrogation and superior understanding.
- Boucheix, J.-M., Lowe, R.K., Groff, J., Paire-Ficout, L., Argon, S., Saby, L., & Alauzet, A. (August, 2011). *Does animation facilitate comprehension of public information graphics? Evidence from eye tracking*. Paper presented at the 14th European Association for Learning and Instruction Conference, Essex, UK.
- Lowe, R. K., & Boucheix, J. M. (2011). Cueing complex animations: Does direction of attention foster learning processes? *Learning and Instruction, 21*, 650-663. doi: 10.1016/j.learninstruc.2011.02.002
- Lowe, R.K., & Keehner, M. (August, 2010). *Comparing search in tactile and visual graphics*. Paper presented at the Comprehension of Text and Graphics Conference, Tübingen, Germany.
- Lowe, R. K., Schnotz, W., & Rasch, T. (2010). Aligning affordances of graphics with learning task requirements. *Applied Cognitive Psychology*, DOI: 10.1002/acp.1712