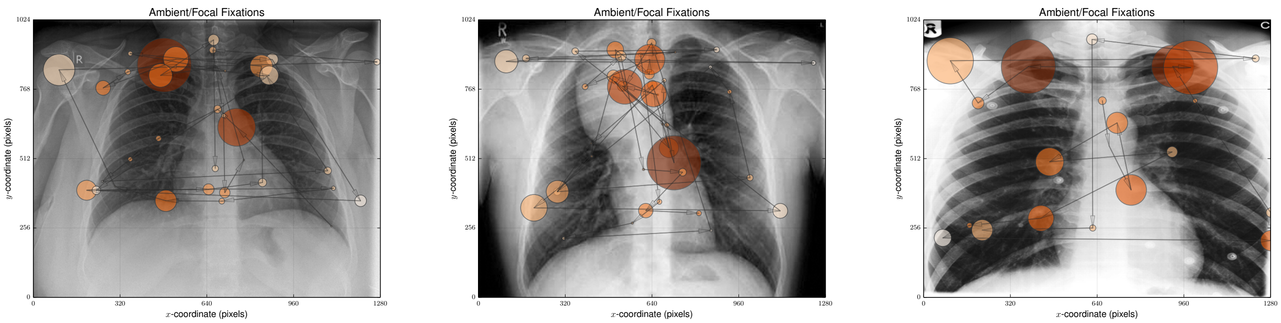
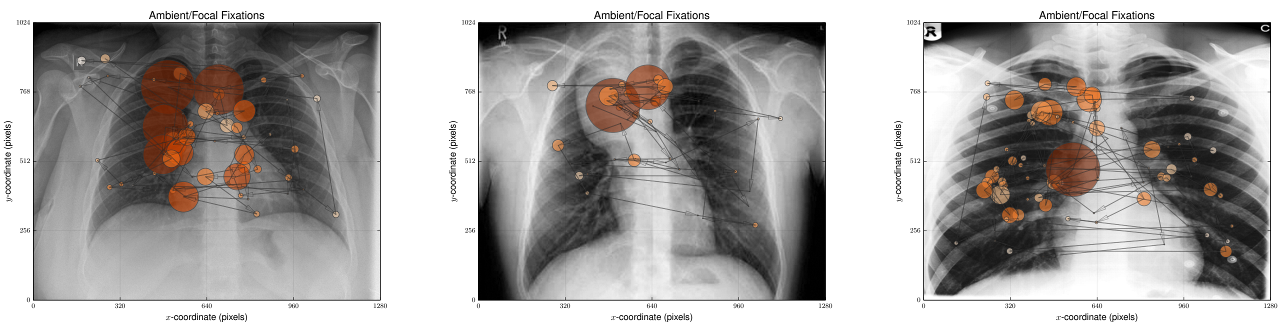
Visualizing Dynamic Ambient/Focal Attention with Coefficient $\mathscr{K}$

A. T. Duchowski *Member, IEEE*, and K. Krejtz



(a) *Expert reading CXR images: normal (left) and abnormal (middle: ant. mediastinal mass, right: apical pneumothorax).*



(b) *Novice reading CXR images: normal (left) and abnormal (middle: ant. mediastinal mass, right: apical pneumothorax).*

Fig. 1. Example of expert/novice scanpaths over Chest X-Ray (CXR) film. CXR images in the middle column feature an anterior mediastinal mass found at approx. $(635, 768)$. Images in the right column feature an apical pneumothorax at approx. $(650, 510)$. Experts tend to execute the visual inspection task considerably faster than novices, with novices tending to dwell longer over abnormalities, if any. Ambient/focal fixation visualization shows a greater preponderance of experts allocating ambient (lighter) fixations in peripheral image regions.

Abstract—Using coefficient $\mathscr{K}$, defined on a parametric scale, derived from processing a traditionally eye-tracked time course of eye movements, we propose a straightforward method of visualizing ambient/focal fixations. The $\mathscr{K}$ coefficient indicates the difference of fixation duration and following saccade amplitude expressed in standard deviation units, facilitating parametric statistical testing. Positive and negative ordinates of $\mathscr{K}$ indicate *focal* or *ambient* fixations, respectively, and are colored by luminance variation depicting relative intensity of focal fixation.

Index Terms—ambient/focal attention, scanpath visualization, eye tracking.

1 Introduction

Visualization plays an increasingly important role in eye tracking analysis. In their EuroVis state-of-the-art (STAR) report, Blascheck et al. [2] review and classify visualization techniques for eye movement data into three categories: point-based, Area-Of-Interest (AOI-based), and those using both. They further distinguish between animated and static, 2D and 3D, in-context and not in-context, as well as interactive and non-interactive visualizations. Finally, visualization techniques are classified as either temporal, spatial, or spatio-temporal.

In this paper we propose a novel (point-based, static, 2D, in-context, spatio-temporal) scanpath visualization of fixations via color mapping between ambient and focal fixations. Colorization of fixations into ambient/focal preserves the traditional spatio-temporal characteristics of scanpath visualizations by conveying order of fixations and fixation durations. However, ambient/focal colorization introduces a novel form of visualization of the dynamic interplay between the focal and ambient modes of visual information processing. Generally, at early stages of scene perception, shorter fixations and longer saccades appear to govern initial scene exploration. Once a target has been identified, longer fixations are followed by shorter saccades suggesting a change to a focal mode of processing [7, 18].

The dynamic pattern of visual attention can be attributed to two modes of acquiring information: exploration and inspection. Pannasch

- *Andrew T. Duchowski is with Clemson University. E-mail: duchowski@clemson.edu.*
- *Krzysztof Krejtz is with the Laboratory of Interactive Technologies, National Information Processing Institute, 00-680 Warsaw, Poland, and the Department of Psychology, University of Social Sci. and Humanities, 03-815 Warsaw, Poland. E-mail: kkrejtz@swps.edu.pl.*

et al. [13] showed a systematic increase in the durations of fixations and a decrease of saccadic amplitudes over the time course of scene perception. This relationship was very stable across the variety of studied conditions, including repeated presentation of similar stimuli, object density, emotional stimuli, and mood induction. In their work, the time courses of fixation durations and saccadic amplitudes were considered as two independent streams of data. We combine both streams into a single dynamic stream defined on a novel parametric scale capturing explicitly the interplay of ambient and focal modes.

2 Background

Blascheck et al. [2] review the state-of-the-art in scanpath visualizations. They note that for a typical scanpath visualization, each fixation is indicated by a circle (or disk), where the radius corresponds to the fixation duration. Saccades between fixations are represented by connecting lines between these circles. The connecting lines may include arrowheads and the fixation circles/disks may include numbers to indicate scanpath order.

Currently, most scanpath visualizations use a constant color to represent fixations. The color may change from scanpath to scanpath, when distinguishing between several individuals if more than one scanpath are composited, but the color does not usually change from fixation to fixation. Our visualization technique exploits this static choice of color and adjusts it at each fixation depending on where the fixation falls on the ambient/focal parametric scale.

Velichkovsky et al. [18] originally suggested characterization of fixations as focal or ambient based on their durations and the amplitude of successive saccades. However, visualizations related to the ambient/focal distinction were limited to graphs resembling histograms depicting either fixation duration or saccade amplitude as a function of viewing time (in 500 ms bins) or as saccade amplitude as a function of fixation duration (in 20 ms bins) [17]. It is important to note that these visualizations are meant to depict the distribution of saccade amplitudes and fixation durations from which one can see that ambient/focal fixations occurred some time during the course of viewing, but not when.

Using the ambient/focal fixation distinction, Follet et al.'s [6] visualizations show the probability of occurrence of fixation type during the time course of viewing. From their visualizations one can see that, for example, ambient fixations are more likely to occur early in the viewing process than focal fixations, but not where.

Krejtz et al. [8] used an ambient/focal attention coefficient, defined as the relation between the current fixation duration and the subsequent saccade amplitude, but did not provide its derivation (see also Biele et al. [1]). The ambient/focal attention coefficient proposed by Krejtz et al. [9] as $\mathcal{K}_i$ transforms both fixation and saccade amplitudes into a standard score (z-score), allowing computation of a focal/ambient attentional coefficient per fixation (and in the aggregate per individual scanpath). Krejtz et al. [10] used the coefficient to analyze map viewing, unfortunately, visualization of the coefficient was never discussed by Krejtz et al. in any of their previous publications.

The coefficient $\mathcal{K}_i$ is calculated for each fixation as the difference between standardized values (z-scores) of the successive saccade amplitude (a_{i+1}) and the current i^{th} fixation duration (d_i) [9]:

$$\mathcal{K}_i = \frac{d_i - \mu_d}{\sigma_d} - \frac{a_{i+1} - \mu_a}{\sigma_a}, \quad i \in [1, n-1] \qquad (1)$$

where μ_d, μ_a are the mean fixation duration and saccade amplitude, respectively, and σ_d, σ_a are the fixation duration and saccade amplitude standard deviations, respectively, computed over all n fixations and hence n $\mathcal{K}_i$ coefficients (i.e., over the entire duration of the scanpath). Coefficient $\mathcal{K}_n$ takes on the value of $\mathcal{K}_{n-1}$ since there is no fixation at $n+1$ from which to compute $\mathcal{K}_n$.

Positive values of $\mathcal{K}_i$ show that relatively long fixations were followed by short saccade amplitudes, indicating focal processing. Analogously, negative values of $\mathcal{K}_i$ refer to the situation when relatively short fixations were followed by relatively long saccades, suggesting ambient processing.

For visualization purposes, the n $\mathcal{K}_i$ coefficients, each associated with the i^{th} fixation, are normalized to facilitate colorization. Subsequently, a color map needs to be selected to produce pleasing and informative visualizations. Because $\mathcal{K}_i$ is associated with each i^{th} fixation, colorization of the fixations depicts when ambient and focal fixations tend to occur and where (per individual scanpath). This visualization, based on the traditional scanpath, depicts the dynamics of ambient/focal attention spatio-temporally, at the expense of depicting statistical trends (e.g., as are possible via histogram-like graphs that lack either spatial or temporal information).

Apart from the novelty of the ambient/focal parametric scale itself, because the visualization technique mainly relies on a suitable choice of color palette, here we briefly only touch on what are likely appropriate color mapping selections.

The rainbow color map is the predominant choice for aggregate gaze visualization (e.g., for heat maps) although it is considered harmful because it [3]:

1. confuses viewers through its lack of perceptual ordering,
2. obscures data via uncontrolled luminance variation, and
3. actively misleads interpretation through the introduction of non-data-dependent gradients.

In essence, the rainbow color map can introduce artificial boundaries in its representation. Ratwani et al. [14] show that the boundaries between red, yellow, green, and blue hues form "visual clusters" that serve as object-like units that can influence reasoning about the graph during cognitive integration. Coincidentally, they demonstrated the importance of these visual cluster boundaries empirically by recording fixations at these boundaries. They state that spectral (rainbow) color palettes should be used but only when differentiation between colors is desired. For gaze visualization, this is a key point, because it suggests the appropriateness of the rainbow color map but largely for *discrimination*, or identification, tasks. For *relative judgements*, Breslow et al. [4] make a compelling argument against the rainbow color map, advocating instead color maps based on brightness (luminance) scales.

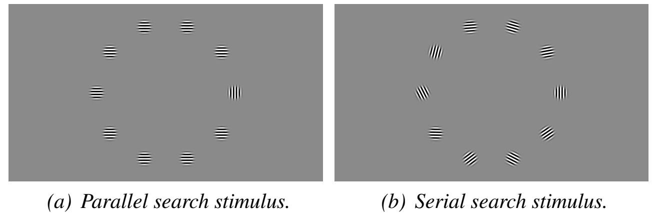
Because our scanpath visualization relies on a continuous parametric scale, colorization via a spectral color palette unnecessarily transforms the scanpath into a visualization meant for identification of regions instead of one showing regions of relative magnitude.

The Python implementation of $\mathcal{K}$ visualization is not tied to any particular color map and allows selection from a variety of convenient choices. There are numerous single, dual, and multi-hue alternatives to the rainbow color map, including palettes from the Colorbrewer website [5].

3 Empirical Validation

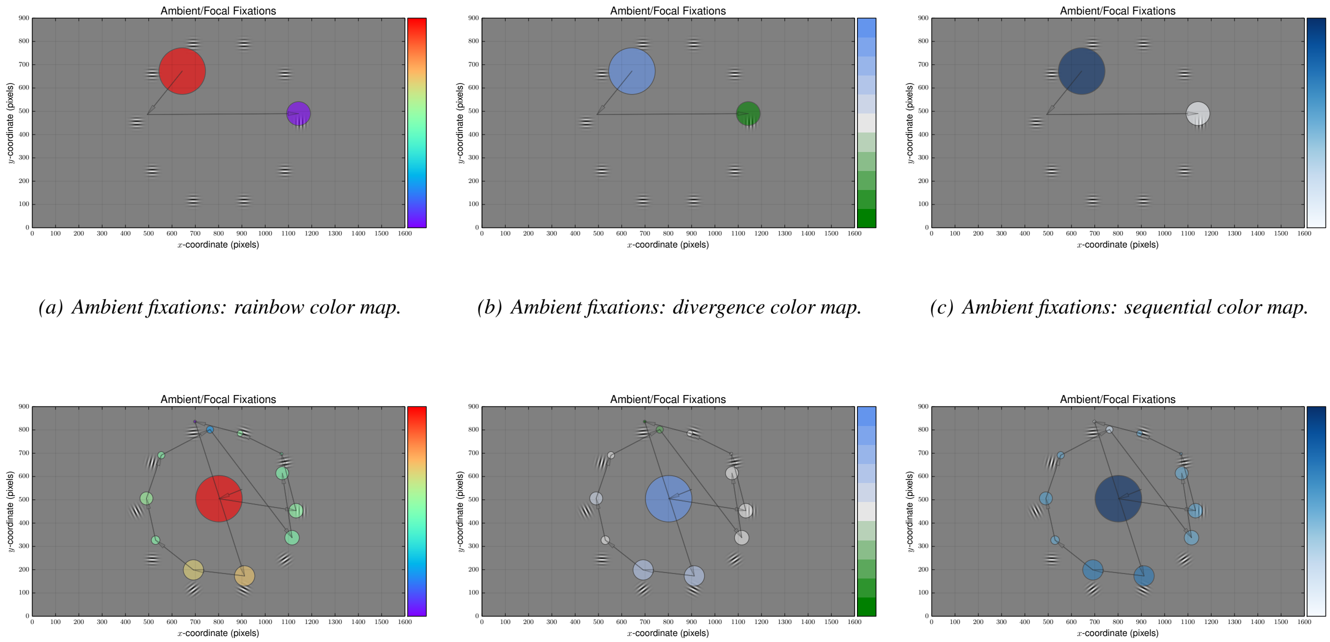
To test different visualization color maps for $\mathcal{K}$, we used data from an experiment designed to replicate empirical procedures reported by Nothdurft [12]. Nothdurft's study showed that serial visual search largely relies on sequential shifts of focal attention whereas no such shifts occur during parallel search.

When performing serial search, therefore, more focal fixations are expected than during parallel search. Due to the pop-out effect during parallel search, fast localization of the target should yield a long saccade (large amplitude) directed to the target. Reaction times reported



(a) Parallel search stimulus. *(b) Serial search stimulus.*

Fig. 2. Visual search stimulus.



(a) Ambient fixations: rainbow color map. *(b) Ambient fixations: divergence color map.* *(c) Ambient fixations: sequential color map.*

(d) Focal fixations: rainbow color map. *(e) Focal fixations: divergence color map.* *(f) Focal fixations: sequential color map.*

Fig. 3. Scanpath composed of mainly ambient (a)–(c) or focal (d)–(f) fixations shown in different choices of color maps.

by Nothdurft support this supposition. They are reminiscent of Treisman and Gelade's response times during *disjunctive* search [16]. The coefficient $\mathscr{K}$ should characterize these eye movements as ambient. Conversely, assuming serial search is composed of longer fixations followed by shorter saccades, $\mathscr{K}$ should reflect serial search as focal.

To replicate Nothdurft's experiment, we followed a within-subjects 2×2 factorial design, with two independent variables of search condition (serial vs. parallel) and target presence (hit vs. rejection). Each participant's task was to find a vertically oriented Gabor patch (the target) within a ring of nine distractor Gabor patches. The distractor patches were either all horizontally oriented (eliciting parallel search) or oriented randomly (eliciting serial search). Examples of stimuli are shown in Figure 2, enhanced for visibility via automatic color balance. Actual stimulus images were not enhanced in this way (see Figure 3). In both examples, the target is present at 3 o'clock.

For brevity we omit details of empirical results and focus on the visualization of $\mathscr{K}$ during searches where the target was present. Results of the experiment are described in detail elsewhere [9].

3.1 Color Map Selection

Analysis bears out the existence of expected types of ambient and focal fixations. Most of the recorded data shows scanpaths that begin with a first focal fixation. This is due to all participants starting by looking at a central fixation point, as per the experimental protocol. The stimulus field appeared after a short delay. During parallel search (see Figures 3(a)–3(c)), participants usually made one large saccade to the intended target, as expected due to the pop-out effect. During serial search (see Figures 3(d)–3(f)), participants usually picked some random Gabor patch and then proceeded to serially inspect the ring of Gabor patches in either clockwise or counter-clockwise order. Figure 3 shows two scanpaths recorded from one participant that is representative of the data set. The scanpaths are colored with three different color maps.

Although the rainbow color map is known to be ineffective, it is nevertheless pervasive, especially in eye tracking visualizations of heat maps. We use the traditional rainbow color map to test the rainbow map's propensity for distinguishing ambient fixations from focal. Figures 3(a) and 3(d) depict fixations during parallel and serial search, respectively. While the color map depicts a visual difference between $\mathscr{K}$ of focal and ambient fixations during serial search (Figure 3(d)), the color hues that were drawn for depicting the focal and (very) ambient fixation during parallel search (Figure 3(a)) do not adequately convey the semantic distance between focal and ambient fixations—why should focal fixations be red and ambient ones dark purple?

In their very effective composition of luminance- and chrominance-based divergent color map, Rogowitz and Lloyd [15] show how luminance can be used to depict magnitude (e.g., terrain elevation) and semantic meaning by splitting the color map (e.g., at sea level). Their choice of color map effectively shows increasing luminance with terrain elevation, with landmass (green) demarcated from water (blue). In our case, a similar argument can be made: $\mathscr{K}$ magnitude is continuous, but there is also a demarcation between focal and ambient fixations at $\mathscr{K} = 0$. To implement this type of color map, we use two different color maps, blue for focal fixations ($\mathscr{K} > 0$) and green for ambient fixations ($\mathscr{K} < 0$). Figures 3(b) and 3(e) show the effect of the divergent color map. Although ambient and focal fixations are distinguished in an adequately dyadic manner, it places the burden on the viewer to remember which color depicts which type of fixation: are focal fixations blue or green? (They are blue.)

Unlike Rogowitz and Lloyd's terrain visualization, which itself carried semantic information (e.g., easily recognizable map of the U.S.), scanpaths do not inherently carry meaning. That is, to use Blascheck et al.'s [2] terminology, they are not in-context unless drawn overlaid atop the stimulus, which, in turn, may not necessarily suggest an inherent viewing order. As such, without any prior expectation as to where or when ambient or focal fixations are expected, specifying divergent colors for their depiction leads to a visual ambiguity.

Our final choice of color map is motivated by visual clarity as well as convenience. Instead of specifying a custom color map for every data set, we would rather just select an appropriate color map from the bevy of choices available in Python's `matplotlib`. The most intuitive choice, based on luminance scaling, is the class of sequential color maps. Figures 3(c) and 3(f) shows the effect of the sequential (blue) color map. Instead of associating a color with ambient or focal type of fixation, lightness of hue indicates intensity of the fixation: darker hues suggest more intense (focal) viewing. Figure 3(c) clearly shows an ambient type of fixation following the initial central focal fixation. In contrast, Figure 3(f) shows relatively darker (more focal) fixations proceeding sequentially along distractor patches until the intended target is fixated.

Figure 1 shows the selection of a similar, sequential color map in orange hues, used to visually distinguish visual inspection of Chest X-Ray (CXR) images as viewed by experts and novices. Radiologists employ a partially endogenous, cognitive visual inspection strategy, related to top-down mechanisms that are based on prior expectations [11], which in turn are couched in training and experience. In the specific case of CXR reading, this strategy may be typified by the ABCDEFGHI mnemonic [19]. The ABCDEFGHI mnemonic guides trainees and practitioners through a series of checks and assessments to inspect **A**irway, **B**ones, **C**ardiac silhouette, **D**iaphragms, **E**xternal soft tissues, **F**ields of the lungs, **G**astric bubble, **H**ila, and **I**nstrumentation. Figure 1 illustrates qualitatively the differences in expert and novice visual strategies: the expert executes the inspection quickly, tending to "check off" the ABCDEFGHI elements, not pausing excessively on any particular element. Visualization of $\mathscr{K}$ readily depicts this strategy, especially in the peripheral image regions (e.g., when inspecting bones, diaphragm). Conversely, the novice tends to dwell longer on each of the elements, often revisiting previously examined regions of the film. An "outside-in" ambient-to-focal strategy is thus not as clearly depicted as it is for the expert.

4 Conclusion

We have presented a visualization of Krejtz et al.'s [9] $\mathscr{K}$ depicting the dynamic interplay between ambient and focal fixations. Visualization is straightforward, resulting from normalization of $\mathscr{K}$ followed by selection of appropriate color map. Although a divergent color map may seem appropriate, we advocate the use of a sequential map instead.

Acknowledgments

We thank Dr. Helena Duchowska (MD, retired) for her help in reading the CXR images and pinpointing the anomalies contained therein.

References

[1] C. Biele, A. Kopacz, and K. Krejtz. Shall we care about the user's feelings?: Influence of affect and engagement on visual attention. In *Proceedings of the International Conference on Multimedia, Interaction, Design and Innovation*, MIDI '13, pages 7:1–7:8, New York, NY, USA, 2013. ACM.

[2] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. Start-of-the-Art of Visualization for Eye Tracking Data. In Borgo, R. and Maciejewski, R. and Viola, I., editor, *EuroGraphics Conference on Visualization (EuroVis)*. EuroVis STAR—State of the Art Report, 2014.

[3] D. Borland and R. M. Taylor II. Rainbow Color Map (Still) Considered Harmful. *IEEE Computer Graphics and Applications*, 27(2):14–17, March/April 2007.

[4] L. A. Breslow, R. M. Ratwani, and J. G. Trafton. Cognitive Models of the Influence of Color Scale on Data Visualization Tasks. *Human Factors*, 51(3):321–338, 2009.

[5] C. Brewer, M. Harrower, A. Woodruff, and D. Heyman. Colorbrewer 2.0: Color advice for maps. Online Resource, Winter 2009. URL: `http://colorbrewer2.org` (last accessed Dec. 2010).

[6] B. Follet, O. Le Meur, and T. Baccino. New insights on ambient and focal visual fixations using an automatic classification algorithm. *i-Perception*, 2(6):592–610, 2011.

[7] D. E. Irwin and G. J. Zelinsky. Eye movements and scene perception: memory for things observed. *Perception and Psychophysics*, 64:882–895, 2002.

[8] I. Krejtz, A. Szarkowska, K. Krejtz, A. Walczak, and A. Duchowski. Audio description as an aural guide of children's visual attention: evidence from an eye-tracking study. In *ETRA '12: Proceedings of the 2012 Symposium on Eye Tracking Research & Applications*, ETRA '12, pages 99–106, New York, NY, USA, 2012. ACM.

[9] K. Krejtz, A. Duchowski, I. Krejtz, A. Szarkowska, and A. Kopacz. Discerning Ambient/Focal Attention with Coefficient $\mathscr{K}$. *Transactions on Applied Perception*, 2015. To appear.

[10] K. Krejtz, A. T. Duchowski, and A. Çöltekin. High-Level Gaze Metrics From Map Viewing: Charting Ambient/Focal Visual Attention. In P. Kiefer, I. Giannopoulos, M. Raubal, and A. Krüger, editors, *Proceedings of the 2nd International Workshop on Eye Tracking for Spatial Research (ET4S)*, September 23 2014.

[11] C. Mello-Thoms, C. F. Nodine, and H. L. Kundel. What attracts the eye to the location of missed and reported breast cancers? In *ETRA '02: Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, pages 111–117, New York, NY, 2002. ACM.

[12] H.-C. Nothdurft. Focal attention in visual search. *Vision Research*, 39:2305–2310, 1999.

[13] S. Pannasch, J. R. Helmert, K. Roth, A.-K. Herbold, and H. Walter. Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, 2(2):1–19, 2008.

[14] R. M. Ratwani, J. G. Trafton, and D. A. Boehm-Davis. Thinking Graphically: Connecting Vision and Cognition During Graph Comprehension. *Journal of Experimental Psychology: Applied*, 14(1):36–49, 2008.

[15] B. Rogowitz and L. Treinish. Data Visualization: The End of the Rainbow. *IEEE Spectrum*, 35(12):52–59, December 1998.

[16] A. Treisman and G. Gelade. A Feature Integration Theory of Attention. *Cognitive Psychology*, 12:97–136, 1980.

[17] P. J. A. Unema, S. Pannasch, M. Joos, and B. Velichkovsky. Time course of information processing during scene perception. *Visual Cognition*, (12(3)):473–494, 2005.

[18] B. M. Velichkovsky, M. Joos, J. R. Helmert, and S. Pannasch. Two Visual Systems and Their Eye Movements: Evidence from Static and Dynamic Scene Perception. In *CogSci 2005: Proceedings of the XXVII Conference of the Cognitive Science Society*, pages 2283–2288, Stresa, Italy, July 2005.

[19] S. A. Vitak, J. E. Ingram, A. T. Duchowski, S. Ellis, and A. K. Gramopadhye. Gaze-Augmented Think-Aloud as an Aid to Learning. In *Proceedings of the SIGCHI Conference on Human Factors in computing systems*, CHI '12, pages 1253–1262, New York, NY, 2012. ACM.