Exploring pupil size variation as a cognitive load indicator in visualization studies

Ugo Molteni, Arzu Çöltekin*

Institute of Interactive Technologies, University of Applied Sciences and Arts Northwestern Switzerland (FHNW)
{ugo.molteni, arzu.coltekin@fhnw.ch)

Pupil size is an elusive eye movement metric in visualization studies as it is sensitive to both internal (cognitive load, task difficulty, medical conditions) and external stimulation (variation in light, i.e., color, contrast). While pupil size has been utilized and critiqued in earlier studies as a metric of cognitive load; systematic examinations of pupil size variation across multiple experiments, involving different visualization types and different levels of expertise, are rare. In this study, to contribute towards a better understanding of pupil size as a cognitive load indicator in visualization studies, we present a preliminary exploration in which we examine if and how pupil size responds to varying visual *stimuli* (different map interfaces, mono and stereo visualizations), and *expertise* levels in two experiments (n=30, n=33). The descriptive statistics suggest, as in some previous studies, that both stimuli and expertise might moderate pupil size variation, and these two factors might interact with each other. Furthermore, we compare right- and left pupil sizes, which appear to differ, especially for mono vs. stereo visualizations. Further analyses are planned to verify the preliminary analyses presented in this paper.

CCS CONCEPTS • H.5: Information interfaces and presentation • Human computer interaction (HCI) • Human centered computing • Visualization • Empirical Studies in Visualization

Additional Keywords and Phrases: Eye tracking, map reading, pupil size, perception, cognition, vision, stereo

1 INTRODUCTION

Since pupil size varies in response to internal cognitive processes (focus, cognitive load, increased attention), physiological conditions (medication, health conditions) *and* external perceptual factors such as light, color and contrast, it is a 'tricky' metric to use and characterize as an indicator of cognitive load, especially in evaluating visualizations which may differ in brightness, color, contrast, and overall visual complexity [e.g., 7, 12, 14]. Despite this ambiguity in visualization studies on why pupil size changes, there is ample evidence since the 1960s that pupils dilate during cognitively complex tasks [9, 10], including certain visuospatial tasks in the context of map use [e.g., 11]. A comprehensive review of this complex relationship between pupil size and task difficulty has been presented by van der Wel & van Steenbergen (2018) in a recent publication [17]. Interestingly, pupil sizes seem to differ between left and right eyes, sometimes due to a medical condition called *anisocoria*, which is observed in approximately 19% of the population [13]. Anisocoria is rarely examined in connection to cognitive load. In a recent study, Çağılltay & Dalveren (2019) reported that in a group of right-handed surgeons, the right pupil was frequently larger than left while executing right-handed or both-

* Corresponding author.

EMICS '21, May 14, 2021, Yokohama, Japan

handed tasks, but not when executing left-handed tasks [2]. While there is limited work that takes pupil size differences between left and right eyes into account, some researchers linked it with attentional processing and experience levels [9, 10]. In visualizations, especially in the case of anaglyph stereoscopic visualizations, having different pupil sizes for left and right eyes can be potentially meaningful also from the perspective that the two eyes are presented with slightly different visual stimuli. Given the above, in this short paper, we explore how pupil sizes vary in response to varying visual complexity of visual stimuli, specifically, two different interface designs and monoscopic vs. stereoscopic viewing of maps. In addition, we examine pupil size variation in connection to expertise, and extend our study also briefly to the differences in left and right pupil sizes.

2 METHODS

The data was collected in two experiments (Exp.1 and Exp.2, n=30 and n=33) with geospatial displays as stimuli, counterbalanced for gender and expertise (i.e., education and professional experience in geospatial domains). In **Exp. 1**, based on a mixed factorial design, two interactive web applications of a national atlas were comparatively evaluated in a randomized controlled laboratory study with the goal to examine their design from a usability perspective. The two applications contain the same information, but differ in their interface and map designs. Stimulus one is called "*Natlas*" (after National Atlas of the United States) whereas the second "*Carto.net*" (after a student project which was temporarily presented under this URL). Participants solved three tasks varying in degree of complexity (not reported in this paper), and data was collected under controlled conditions using a screen-based eye tracker (Tobii TX300). Task and stimuli order were rotated using a Latin-squares approach. Participant performance based on response accuracy and time, usability outcomes, and sequence analysis of participants' eye movements have been published from Exp.1 [1, 3]. **Exp.2** is also a mixed factorial design, and it compares display types (stereoscopic vs. monoscopic, *i.e., stereo* vs. *mono*) in a perception study. The displays are static and contain satellite imagery of terrain in mono and stereo. The main goal in the study is to compare participants' depth perception in the presence of an optical illusion (where valleys appear as ridges and vice versa) [4, 5, 6, 8]. Participants were asked to identify an indicated landform as "clearly a valley" to "clearly a ridge" using a 5-point Likert scale. Stereo and mono conditions contained the same set of terrains and all participants worked with stereo and mono. Exp.2 was conducted in blocks in rotated order, and the stimuli in each block were randomized. The eye movement data was collected in the same lab with the same eye tracker as in Exp1. In this publication, our focus is on participants' pupil size variation in the studies outlined above. Below, we present preliminary (visual and exploratory) outputs from our pupil size variation analysis, based on Matlab R2020b [18].

3 RESULTS AND DISCUSSION

For Exp.1, we present pupil size variation for geographers (expert users) vs. non-geographers, and for stimulus type in Figure 1. In this plot we show the pupil size variation for participants' left eye only. Right eye pupil variation overall exhibits the same pattern, with some individual differences (elaborated later). Figure 1 suggests that the effect of stimulus on pupil diameter is more apparent than that of expertise (compare columns): Both geographers and non-geographers exhibit overall larger pupil diameters with Carto.net than with Natlas. If we focus on expertise (compare rows), the variation is more subtle, however, the change based on stimulus is more pronounced for geographers than for non-geographers where variation occurs in a narrower window with Natlas than with Carto.net. This potentially suggests an interaction between stimulus complexity and expertise levels. Overall Natlas interface required more steps from the users and participants were considerably faster with carto.net than with Natlas; however, response accuracy was higher with Natlas [1]. Perhaps Natlas interface was *necessarily* complex. The alternative explanation is that carto.net had brighter colors or more contrast, though this interpretation requires an assessment of saliency and further analysis of colors used by the two applications (stimuli examples can be seen in [1, 3]). Next, we present a distribution plot for Exp.2 for the three expertise groups and two stimulus types (Figure 2). As in Exp.1, data is shown only for the left eye. Figure 2 illustrates apparent differences both based on expertise and stimulus in participants' pupil size, and most strikingly, in stereo pupil diameter reaches to 6mm. For the mono condition in Exp.2, *and* in the Exp.1 where stimuli were mono, pupil diameter does not exceed ~4.5mm. Stereo might require more focus, as the eyes are somewhat strained by the mode of operation due to vergence focus conflict; and it may also lead to more attentional processing, because there is simply more space in three-dimensions to explore than there is in mono. A more immediate explanation is perceptual, as anaglyph (red/blue) glasses filter the incoming light to some degree and the larger pupil sizes in stereo condition may be explained possibly mostly by this (less light, larger pupils). The variation among different levels of expertise, especially in the case of the non-geographers in stereo condition (lower right panel in Figure 2) is interesting, as it appears that stereo condition is especially challenging or interesting for non-experts. The fact that expertise appears somewhat relevant speaks for the cognitive load rather than perceptual stimulation, though this speculation needs to be verified in further inferential analyses. What is not shown in Figures 1 and 2 (due to lack of space in this short paper) is that when we plot the pupil sizes for left and right eyes, we see many individual differences (see Figure 3 for an example).

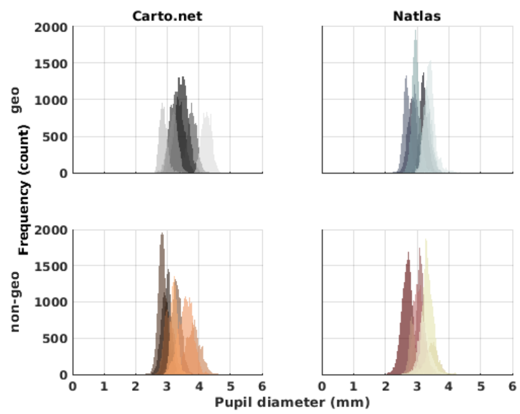

Figure 1: Pupil diameters by stimulus type and two levels of expertise in Exp.1. Each bell curve is a participant's pupil size distribution. Vertical axis shows a count of how frequently pupil diameter is reported in the data set (every ~16ms, i.e., eye tracker's data capture frequency), thus the vertical axis is an implicit expression of experiment duration per participant.

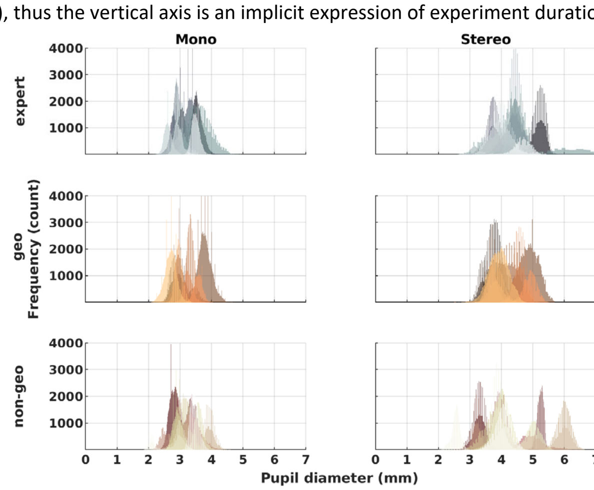

Figure 2: Pupil diameters organized by stimulus type and three levels of expertise in Exp. 2. Again, each bell curve is a participant's pupil size distribution. Here "expert" is someone who works as a researcher specifically with satellite images, "geo" is someone who studied geography or in a closely related field, and "non-geo" is someone with little to no expertise in geospatial fields.

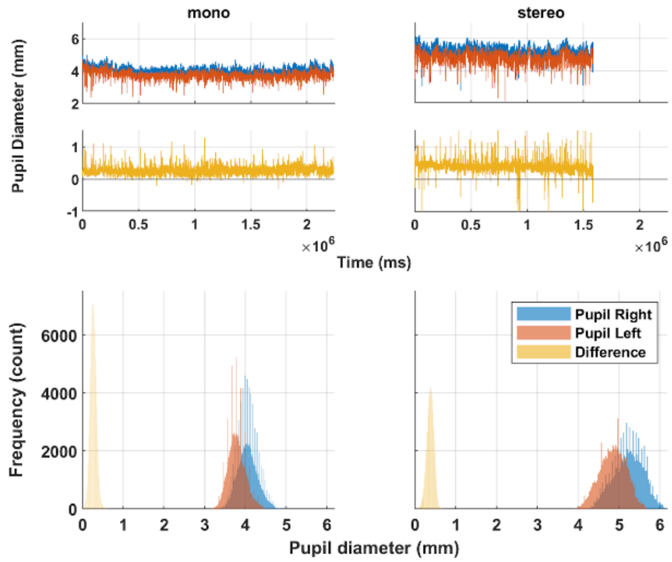

Figure 3: Pupil diameter of an individual participant viewing the mono (left) vs. stereo (right) images in Exp.2. The top part (both left and right) shows the time series plot, and in the lower part we see the difference (right minus left pupil diameter) on the left side of each panel, then pupil size distribution of the participant for the right and left eyes on the right side of each panel.

While Figure 3 illustrates one participant, based on a simple count, we found visible differences between left and right pupil diameter in ~39% of our participants across the two experiments (considerably higher than the ~19% reported [13]). This difference is more pronounced in the stereo condition (on average -0.006mm for mono, 0.013mm for stereo) in Exp.2. In Exp.1, both participant groups have notably larger left pupils when solving the tasks with Carto.net (geo: -0.06, non-geo: -0.05mm) than with Natlas (geo: 0.02, non-geo: 0.01mm). Overall, right-left pupil size differences are more apparently linked with stimulus (especially in the case of mono vs. stereo) than with expertise, however in Exp.1, the right-left gap is more pronounced for geo group with Carto.net.

4 CONCLUSIONS AND OUTLOOK

As mentioned before, pupil size differences during visual stimulus use could be due to perceptual reasons or visual and/or task complexity (consequently, an expression of cognitive load). On the other hand, any differences in pupil size based on expertise or task complexity could be an expression of cognitive load. In this preliminary analysis based on descriptive statistics we see that in some cases experts and non-experts' pupil sizes differ, possibly moderated by display type. Furthermore, there is a considerable gap in pupil size variation for right and left eyes of the participants, implications of which need to be further explored. Since we did not employ any inferential analysis in this short paper, our observations should be taken only as hypotheses at this point. To confirm these preliminary observations, and to better understand the reasons behind them, we plan to analyze pupil size variations more in depth with inferential analyses; and examine interactions between display type, task complexity, expertise as well as temporal patterns to control for fatigue. Also as an outlook, we plan to analyze the differences between left and right eye pupil sizes in connection to handedness.

ACKNOWLEDGMENTS

We are grateful to our past collaborators Benedikt Heil, Simone Garlandini and Martina Meyer for their efforts in collecting the eye movement data.

REFERENCES

[1] Julien Biland, and Arzu Çöltekin. 2017. An empirical assessment of the impact of the light direction on the relief inversion effect in shaded relief maps: NNW is better than NW. Cartography and Geographic Information Science 44(4), 358–72. https://doi.org/10.1080/15230406.2016.1185647

[2] Nergiz Ercil Cagiltay and Gonca Gokce Menekse Dalveren. 2019. Are left- and right-eye pupil sizes always equal? Journal of Eye Movement Research 12(2), 1–9. https://doi.org/10.16910/jemr.12.2.1

[3] Arzu Çöltekin, and Julien Biland. 2019. Comparing the terrain reversal effect in satellite images and in shaded relief maps: an examination of the effects of color and texture on 3D shape perception from shading. International Journal of Digital Earth 12(4), 442–59. https://doi.org/10.1080/17538947.2018.1447030

[4] Arzu Çöltekin, Benedikt Heil, Simone Garlandini, and Sara Irina Fabrikant. 2009. Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis. Cartography and Geographic Information Science 36(1), 5–17. https://doi.org/10.1559/152304009787340197

[5] Arzu Çöltekin, Sara I. Fabrikant, and Martin Lacayo. 2010. Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. International Journal of Geographical Information Science 24(10), 1559–75. https://doi.org/10.1080/13658816.2010.511718

[6] Arzu Çöltekin, Gianna Hartung, and Martina Meyer. 2019. Deconstructing the relief inversion effect: Contributors of the problem and its solutions. Abstracts of the ICA 1, 1–2. https://doi.org/10.5194/ica-abs-1-48-2019

[7] Rahul Gavas, Debatri Chatterjee, and Aniruddha Sinha. 2017. Estimation of cognitive load based on the pupil size dilation , 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, pp. 1499–1504. https://doi.org/10.1109/SMC.2017.8122826

[8] Gianna Hartung, and Arzu Çöltekin. 2020. Fixing an illusion – an empirical assessment of correction methods for the terrain reversal effect in satellite images. International Journal of Digital Earth 13(10), 1135–50. https://doi.org/10.1080/17538947.2019.1681526

[9] Eckhard H. Hess, and James M. Polt. 1964. Pupil size in relation to mental activity during simple problem-solving. Science 143(3611), 1190–92. https://doi.org/10.1126/science.143.3611.1190

[10] Daniel Kahneman. 1973. Attention and effort. Englewood Cliffs, NJ. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

[11] Peter Kiefer, Ioannis Giannopoulos, Andrew Duchowski, and Martin Raubal. 2016. Measuring cognitive load for map tasks through pupil diameter. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 323–37. https://doi.org/10.1007/978-3-319-45738-3_21

[12] Krzysztof Krejtz, Andrew T. Duchowski, Anna Niedzielska, Cezary Biele, and Izabela Krejtz. 2018. Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. Ed. Susana Martinez-Conde. PLOS ONE 13(9), e0203629. https://doi.org/10.1371/journal.pone.0203629

[13] Byron L. Lam, H. Stanley Thompson, and James J. Corbett. 1987. The prevalence of simple anisocoria. American Journal of Ophthalmology 104(1), 69–73. https://doi.org/10.1016/0002-9394(87)90296-0

[14] Oskar Palinko, Andrew L. Kun, Alexander Shyrokov, and Peter Heeman. 2010. Estimating cognitive load using remote eye tracking in a driving simulator. Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10. New York, New York, USA: ACM Press, p. 141. https://doi.org/10.1145/1743666.1743701

[15] William D. Poynter, 2017. Pupil-size asymmetry is a physiologic trait related to gender, attentional function, and personality. Laterality: Asymmetries of Body, Brain and Cognition 22(6), 654–70. https://doi.org/10.1080/1357650X.2016.1268147

[16] Basil Wahn, Daniel P. Ferris, W. David Hairston, and Peter König. 2017. Pupil size asymmetries are modulated by an interaction between attentional load and task experience. BioRxiv. https://doi.org/10.1101/137893

[17] Pauline van der Wel and Henk van Steenbergen. 2018. Pupil dilation as an index of effort in cognitive control tasks: A review. Psychonomic Bulletin & Review 25(6), 2005–15. https://doi.org/10.3758/s13423-018-1432-y

[18] MATLAB. 2021. R2019b Update 8 (9.7.0.1586710). Natick, Massachusetts: The MathWorks Inc.