Defining Areas of Interest for Eye-Tracking Data: Implementing a Systematic Approach

Joanna Coltrin * Eric McKinney † Breanna Studenka ‡ Jürgen Symanzik §

Abstract
There exists a variety of methods that have been used to analyze eye-tracking data. One of the general methods involves the use of Areas of Interest (AOIs). AOIs are predefined areas of an image used to determine characteristics of eye-tracking data. While most AOIs are defined by hand, we discuss the use of systematic AOIs and the application of the systematic Voronoi Tessellation Method. Differentiated eye-tracking data can then be compared within the AOIs to determine whether subjects from a treatment group looked at the images differently than subjects from a control group and where those differences occurred.

Key Words: Voronoi Tesselation Method, USU Posture Study, Data Visualization

1. Introduction

A primary goal of eye-tracking technology and analyses is to gain insight into what individuals give their attention to. To collect this information, the eye-tracking device records eye movements of a participant and produces a series of coordinates that describe the location of the subject's gaze. The raw x and y coordinates the eye-tracking device produces are called gaze points. There are several methods researchers can use to analyze gaze points produced by eye trackers. This article focuses on a specific method which uses Areas of Interest (AOIs).

In any given eye-tracking study, each participant is asked to look at a provided stimulus. This stimulus is referred to as the image of interest. According to Hessels et al. (2016), AOIs are predefined areas of an image that are used to determine characteristics of eye-tracking data. These AOIs divide the image of interest into segments determined by the researcher. While many AOIs are defined by subjective hand-drawn segments (Hessels et al., 2016), these areas could be defined more systematically.

The USU Posture Study provides the opportunity to analyze eye-tracking data using systematic AOIs. This study aims to determine whether judgment of others' action capabilities is based on one's own action experiences (Symanzik et al., 2017, 2018; Studenka et al., 2020). This study uses two groups of participants: those with extensive recent yoga experience serve as the treatment group and those with minimal yoga experience serve as the control group. The primary goal of this article is to analyze differences in gaze point proportions within each AOI between the treatment and the control groups in 22 different human postures.

This article is structured as follows: In Section 2, we provide further details of the USU Posture Study. In Section 3, we introduce the Voronoi Tessellation Method and we outline

*Department of Mathematics and Statistics, Utah State University, Logan, UT 84322–3900, USA. E-mail: jcoltrin@aggiemail.usu.edu
†Department of Mathematics and Statistics, Utah State University, Logan, UT 84322–3900, USA. E-mail: ericmckinney77@gmail.com
‡Department of Kinesiology and Health Science, Utah State University, Logan, UT 84322–7000, USA. E-mail: Breanna.studenka@usu.edu
§Department of Mathematics and Statistics, Utah State University, Logan, UT 84322–3900, USA. E-mail: symanzik@math.usu.edu

Figure 1: Posture IDs 1-22 ordered from left to right, then top to bottom.

the application to creating AOIs for our eye-tracking data in Section 4. Section 5 reports the results of the study. In Section 6, we provide our conclusions and an outlook on the next steps in our ongoing analysis. All of our visualizations and analyses are conducted with the `R` statistical computing platform (R Core Team, 2019).

2. The USU Posture Study

The primary research question for the USU Posture Study asked: ”Does a person’s judgement of action capabilities of another person depend on personal knowledge and experience with human movement?” (Symanzik et al., 2017, 2018; Studenka et al., 2020). Specifically, when subjects look at an image to judge the stability of a posture, are there differences between subjects with and without recent yoga experience? If so, where do those differences

Figure 2: Methods for defining AOIs. Hand-drawn (left), grids (middle), and Voronoi (right).

occur?

To answer these questions, data were collected from 40 subjects. The control group consisted of 20 participants with minimal yoga experience and the treatment group consisted of 20 participants with extensive yoga experience. Each of the 40 participants was asked to look at a series of 22 images of an actor holding a posture in random order. Eye-tracking technology captured the gaze points where each participant looked to make the judgement. Each image was given an identification number referred to as the posture ID. Figure 1 shows each of the 22 images labeled as Posture IDs 1-22 that were viewed by all 40 participants.

3. Systematic AOI Methods

A variety of techniques have been used to create AOIs in the existing literature. The most common method for defining AOIs is having an expert in the application area being studied draw the AOIs by hand (Hessels et al., 2016; Cantoni et al., 2012; Fichtel et al., 2019). Because this approach is highly subjective, this method is far from ideal. Due to the subjective nature of this method, more systematic approaches have been developed. A systematic approach to defining AOIs that is most frequently applied involves using regular grids across the observation area (Wästlund et al., 2018). This method is useful since these AOIs are content independent, and consequently, easily generated. However, inferential statistics have shown to be dependent on the granularity of the grid (Duchowski, 2007; McKinney and Symanzik, 2019).

Another approach to defining AOIs systematically is called the Voronoi method adapted from Voronoi (1909). This method requires the center of each AOI to be defined. Lines that separate each AOI at the exact midpoint from the centers are drawn and form tessellations. These tessellations become the AOIs. Each gaze point or fixation point belongs to the AOI where the distance from that point to the AOI center is the shortest. This approach has the advantage of being a systematic approach that is not being restricted to the shape or size of grid cells. This approach is unique in that every part of the image of interest is defined under an AOI because the outermost AOIs extend to the edges of the image. However, the centers of the AOIs, while optimally defined by content subject experts, can be still be

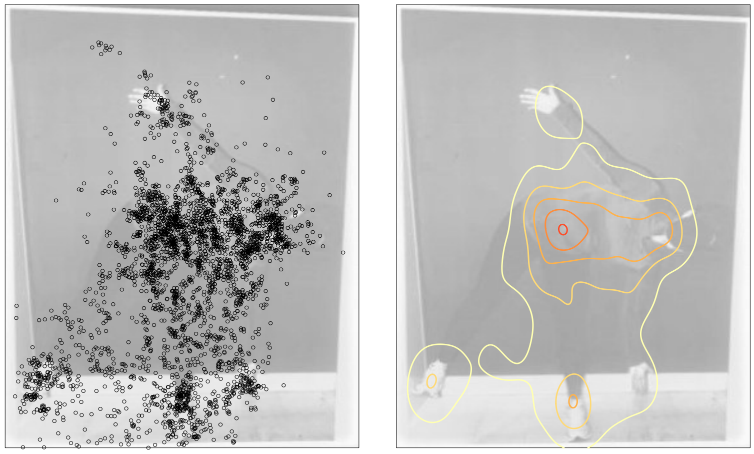

Figure 3: Scatter plot (left) and contour plot (right) of the collective set of gaze points from all subjects for posture with ID 5.

subjective in nature. Nyström et al. (2013) used the Voronoi method as calibration targets. A depiction of these three approaches to defining AOIs can be seen in Figure 2 where the left image shows a "hand-drawn" method, the center image shows the use of a regular grid, and the right image depicts Voronoi tessellations of the AOIs.

4. Voronoi Tessellation AOI Application to the USU Posture Study Data

The overall goal of the USU Posture Study is to compare gaze points from the treatment and the control groups. The goal for this article is to use a systematic method for defining AOIs to analyze the differences in gaze points patterns between the treatment and the control groups. Our hypotheses are as follows: H_0: There are no differences in gaze point proportions within AOIs between the treatment and the control groups. H_A: There are differences in gaze point proportions within AOIs between the treatment and the control groups. To visualize the eye-tracking data, we chose to use scatter plots in instances where individual gaze point locations matter and contour plots to understand where the bulk of the gaze points were concentrated. These visualizations applied to gaze points from all 40 participants looking at posture with ID 5 are represented in Figure 3.

Every participant was shown a series of dots in each corner of an image at the beginning and end of the study to serve as a calibration check. When we combined the data for all 40 participants from the beginning and end of the study, we found that the gaze points were not centered over the calibration dots as depicted in the left and center plots of Figure 4. An affine transformation (Flusser and Suk, 1993) was applied to the gaze points for each participant separately which more accurately plotted the gaze points as is seen in the right plot of Figure 4. An affine transformation was chosen because it preserves colinearity and ratio of distances through any combination of rotation, translation, dilation, and shear transformations (Flusser and Suk, 1993). The Morpho `R` package (Schlager, 2017) was used to compute the transformation matrices and apply the transformation to all x and y coordinates from the gaze points produced by the eye tracker.

The centers of each AOI for each of the 22 postures were defined by a kinesiology expert who identified major joints, the center of mass, and the head. The centers were numbered for consistency, however, not all major joints were visible in every posture im-

Figure 4: Calibration photo (left) contour plot of raw data (middle) and contour plot of transformed data (right).

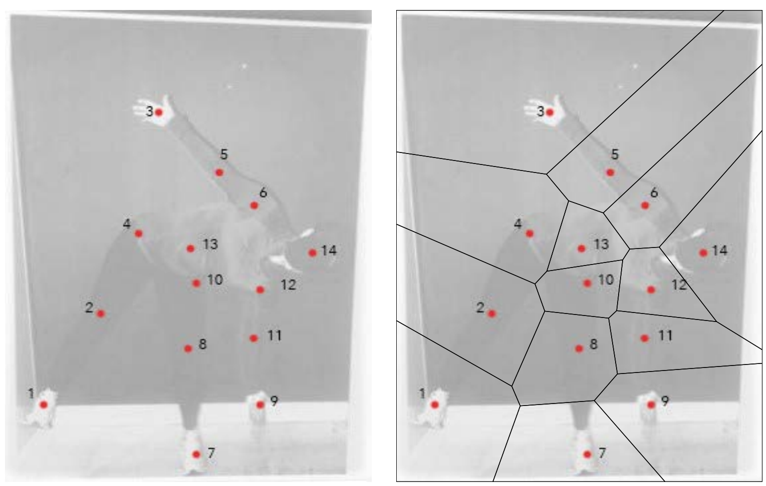

Figure 5: Expert-defined AOI centers for the posture with ID 5 (left) and Voronoi Tessellations for the posture with ID 5 (right).

age. Therefore, if any of these defined AOI centers were obstructed from view, they were excluded from the analysis of that particular posture. The identified centers were used to calculate the tessellations for posture ID 5 as shown in Figure 5.

We used two analyses to test the hypotheses. The Chi-square test for independence (McHugh, 2013) was used to compare the proportion of treatment and control gaze points in each AOI where the degrees of freedom were one less than the number of AOIs in each posture. However, the Chi-square test of independence makes the assumption that 80% of cells (AOIs) have expected values of at least 5 observations (gaze points) and that no cell has an expected value of 0 observations. Hence, the Fisher's exact test (Upton, 1992) was used as a non-parametric alternative to the Chi-square test for independence for any postures that did not meet the assumptions for the Chi-square test for independence. We used simulated p-values for the Fisher's exact test based on 9999 replicates. Both tests were conducted at the 5% significance level.

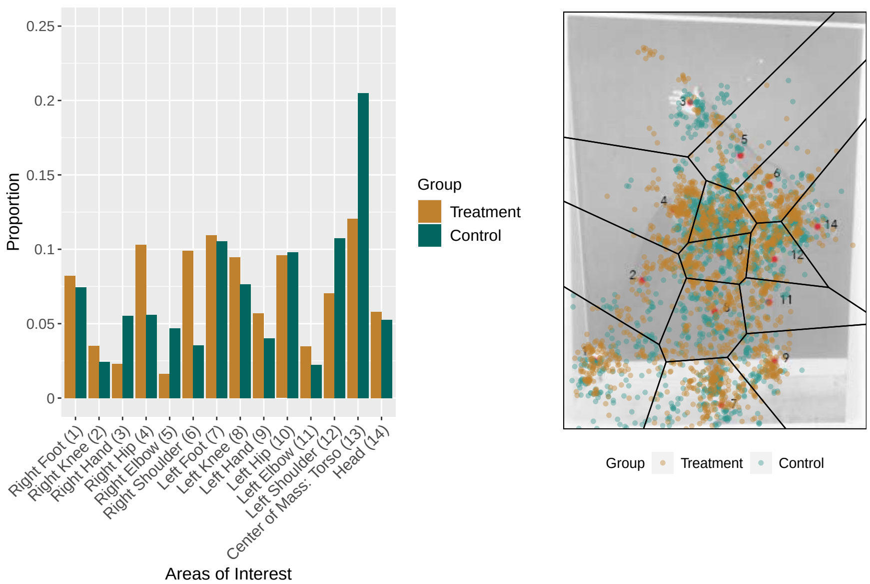

Figure 6: Comparison of proportion of treatment and control gaze points by AOI for the posture with ID 5 (left). Scatterplot of gaze points on the posture with ID 5 with overlaid Voronoi Tessellations (right).

The Bonferroni adjustment (Napierala, 2012) was used as a conservative approach to correct for multiple comparisons for both the Chi-square p-values and the Fisher's exact p-values. The Bonferroni adjustment accounts for the multiple comparisons by multiplying each p-value by the number of tests conducted (Napierala, 2012).

5. Results

The most significant results were obtained from the posture with ID 5 which was introduced in Figures 3 and 5. The analyses on this posture revealed a Chi-square p-value of 3.36e^{-33} and a Bonferroni adjusted Chi-square p-value of 7.39e^{-32}. The Fisher's exact p-value was 0.0001 (1.00e^{-04}) with the Bonferroni adjusted Fisher's p-value of 0.0022 (2.20e^{-03}). Thus, we see that the two tests lead to the conclusion that there is highly significant evidence that there are differences in gaze point proportions within AOIs between the treatment and the control groups for the posture with ID 5.

The statistical test results are confirmed in Figure 6 which explicitly shows how the treatment and control proportions differ in each AOI and depicts where the gaze points are located specifically on the image. The bar chart in Figure 6 shows that the proportion of gaze points from the control group are much higher in the center of mass and left shoulder AOIs than the proportion of gaze points from the treatment group. We also observe that the proportion of gaze points for the treatment group are much higher in the right shoulder and right hip AOIs when compared to the proportion of gaze points from the control group. The p-values for the posture with ID 5 from both tests are highly significant. In fact, this posture has the lowest Chi-squared p-values among all 22 postures which can be seen in Table 1.

In contrast the posture with ID 8 has the highest Chi-squared p-values. However, these p-values also show statistical significance. Figure 7 shows a side-by-side barchart of proportions of gaze points per AOI by treatment group and a scatterplot of the gaze points for the posture with ID 8. Table 1 shows that the posture with ID 8 has a Chi-square p-value

Table 1: P–values for each Posture ID (Alternative = "Two-Sided"). All test outcomes are highly statistically significant, rejecting the null hypothesis in favor of the alternative, i.e. there are differences in gaze point proportions within AOIs between the treatment and the control groups.

Posture ID	Chi-square test for independence	Bonferroni adjusted Chi-square test	Fisher's exact test	Bonferroni adjusted Fisher's exact test
1	$2.11e^{-18}$	$4.64e^{-17}$	$1.00e^{-04}$	$2.20e^{-03}$
2	$1.17e^{-27}$	$2.57e^{-26}$	$1.00e^{-04}$	$2.20e^{-03}$
3	$4.05e^{-27}$	$8.91e^{-26}$	$1.00e^{-04}$	$2.20e^{-03}$
4	$1.45e^{-25}$	$3.19e^{-24}$	$1.00e^{-04}$	$2.20e^{-03}$
5	$3.36e^{-33}$	$7.39e^{-32}$	$1.00e^{-04}$	$2.20e^{-03}$
6	$2.70e^{-18}$	$5.94e^{-17}$	$1.00e^{-04}$	$2.20e^{-03}$
7	$5.59e^{-33}$	$1.23e^{-31}$	$1.00e^{-04}$	$2.20e^{-03}$
8	$3.45e^{-05}$	$7.59e^{-04}$	$1.00e^{-04}$	$2.20e^{-03}$
9	$4.99e^{-12}$	$1.10e^{-10}$	$1.00e^{-04}$	$2.20e^{-03}$
10	$9.34e^{-15}$	$2.05e^{-13}$	$1.00e^{-04}$	$2.20e^{-03}$
11	$3.67e^{-18}$	$8.07e^{-17}$	$1.00e^{-04}$	$2.20e^{-03}$
12	$1.03e^{-20}$	$2.27e^{-19}$	$1.00e^{-04}$	$2.20e^{-03}$
13	$2.59e^{-21}$	$5.70e^{-20}$	$1.00e^{-04}$	$2.20e^{-03}$
14	$5.64e^{-22}$	$1.24e^{-20}$	$1.00e^{-04}$	$2.20e^{-03}$
15	$2.03e^{-12}$	$4.47e^{-11}$	$1.00e^{-04}$	$2.20e^{-03}$
16	$1.24e^{-26}$	$2.73e^{-25}$	$1.00e^{-04}$	$2.20e^{-03}$
17	$4.58e^{-17}$	$1.01e^{-15}$	$1.00e^{-04}$	$2.20e^{-03}$
18	$1.01e^{-14}$	$2.22e^{-13}$	$1.00e^{-04}$	$2.20e^{-03}$
19	$2.79e^{-26}$	$6.14e^{-25}$	$1.00e^{-04}$	$2.20e^{-03}$
20	$4.90e^{-07}$	$1.08e^{-05}$	$1.00e^{-04}$	$2.20e^{-03}$
21	$1.76e^{-31}$	$3.87e^{-30}$	$1.00e^{-04}$	$2.20e^{-03}$
22	$3.10e^{-18}$	$6.82e^{-17}$	$1.00e^{-04}$	$2.20e^{-03}$

of $3.45e^{-05}$ and a Bonferroni adjusted Chi-square p-value of $7.59e^{-04}$. The Fisher's exact test gives a p-value of 0.0001 and a Bonferroni adjusted Fisher's p-value of 0.0022.

Figure 7 clarifies how the treatment and control proportions differ in each AOI and depicts where the gaze points are located specifically on the image. The bar chart in Figure 7 shows that the proportion of gaze points from the treatment group are much higher in the center of mass and left shoulder areas than the proportion of gaze points from the control group. We observe that the proportion of gaze points for the control group are much higher in the left elbow and right hip/ left hand areas when compated to the proportion of gaze points from the treatment group. The center of mass AOI is an AOI in both postures that demonstrates a dramatic difference in proportions of gaze points. However, it is interesting to note that group with the higher proportion of gaze points is not consistent in these two postures.

Overall, Table 1 shows that all Fisher's exact test p-values are the same: 0.0001. These p-values were computed using 9999 permutations. The p-value 0.0001 demonstrates that, assuming there is no difference in the proportions for the treatment and the control groups, none of the permutations distributed the data more disproportionately than the observed data. The consistency in p-values for the Fisher's exact test explains the similar consistency in p-values for the Fisher's Bonferroni adjusted p-values and we see that these p-values of 0.0022 are calculated by multiplying the Fisher's exact p-values by 22, the number of postures analyzed.

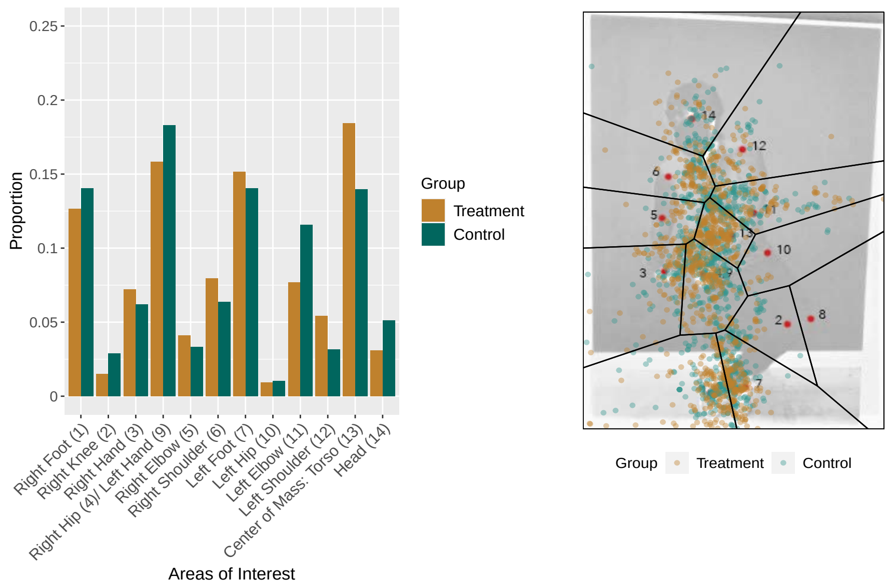

Figure 7: Comparison of proportion of treatment and control gaze points by AOI for the posture with ID 8 (left). Scatterplot of gaze points on the posture with ID 8 with overlaid Voronoi Tessellations (right).

6. Conclusions and Outlook

All p-values provided in Table 1 show that the 22 postures have gaze point proportions that are significantly different between the treatment and the control groups. Therefore, we can conclude that there are differences in gaze point proportions within AOIs between the treatment and the control groups. These results suggest that there is an association between personal knowledge and experience with human movement and a person's judgement of action capabilities of another person.

In this article, we used the Voronoi Tessellation method to define areas of interest and analyze gaze point proportions differences between the treatment and the control groups. While we can conclude that there are differences in gaze point proportions within AOIs between the treatment and the control groups, we also suspect that the results are dependent on the number of AOIs and the distance of the expert-defined AOI centers. We plan to use the information we gather from the margin of error observed around the calibration dots and the gaze points themselves to inform the kinesiologist who will create a second data driven round of AOIs. The results from the second round of AOIs can then be compared to the results presented in this article.

Furthermore, there is a fourth approach to define AOIs that Hessels et al. (2016) explained which could also be implemented in our future analyses. This approach is called the Limited-Radius Voronoi Tessellation (LRVT) method. This method uses the same underlying idea as the original Voronoi method in creating the AOIs with expert-defined AOI centers. However, this modified approach adds a maximum radius from the center that limits the extent of each AOI. The LRVT method provides the same advantages as the Voronoi method and adds the benefit of a user-defined maximum radius from each defined center. Researchers can be more confident when classifying any eye-tracking data found in a particular AOI that outlying gaze points are not inappropriately classified.

The use of the LRVT method to define AOIs merits its future application in the USU Posture Study. We anticipate applying this method using knowledge gained from the margin of error found from the calibration gaze points to inform the chosen maximum radius.

Acknowledgements

We would like to thank our colleagues and fellow students at USU, particularly Dr. Brennan Bean, Dr. Todd Moon, Kristina Casos, Melanie Athens, and Madison Hansen, for helpful discussions and for their efforts contributing to the USU Posture Study (Symanzik et al., 2017, 2018; Studenka et al., 2020).

References

Cantoni, V., Porta, M., De Maio, L., Distasi, R., Nappi, M., 2012. Towards a Novel Technique for Identification Based on Eye Tracking. In: Workshop on Biometric Measurements and Systems for Security and Medical Applications Proceedings. IEEE, Salerno, Italy, pp. 1–4.

Duchowski, A. T., 2007. Eye Tracking Methodology: Theory and Practice. Springer, New York, NY.

Fichtel, E., Lau, N., Park, J., Parker, S. H., Ponnala, S., Fitzgibbons, S., Safford, S. D., 2019. Eye Tracking in Surgical Education: Gaze-Based Dynamic Area of Interest Can Discriminate Adverse Events and Expertise. Surgical Endoscopy 33 (7), 2249–2256.

Flusser, J., Suk, T., 1993. Pattern Recognition by Affine Moment Invariants. Pattern Recognition 26 (1), 167–174.

Hessels, R. S., Kemner, C., van den Boomen, C., Hooge, I. T., 2016. The Area-of-Interest Problem in Eyetracking Research: A Noise-Robust Solution For Face and Sparse Stimuli. Behavior Research Methods 48 (4), 1694–1712.

McHugh, M. L., 2013. The Chi-Square Test of Independence. Biochemia Medica 23 (2), 143–149.

McKinney, E., Symanzik, J., 2019. Modifications of the Syrjala Test for Testing Spatial Distribution Differences Between Two Populations. In: 2019 JSM Proceedings. American Statistical Association, Alexandria, VA, pp. 2518–2530.

Napierala, M. A., 2012. What is the Bonferroni Correction? American Academy of Orthopaedic Surgeons Now 6 (4), 40.

Nyström, M., Andersson, R., Holmqvist, K., Van De Weijer, J., 2013. The Influence of Calibration Method and Eye Physiology on Eyetracking Data Quality. Behavior Research Methods 45 (1), 272–288.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL `https://www.R-project.org/`

Schlager, S., 2017. Morpho and Rvcg – Shape Analysis in R. In: Zheng, G., Li, S., Szekely, G. (Eds.), Statistical Shape and Deformation Analysis. Academic Press, London, UK, pp. 217–256.

Studenka, B. E., Athens, M., Casos, K., Coltrin, J., McKinney, E., Symanzik, J., June 2020. The Influence of Postural Stability and Yoga Experience on Perceptions of Other's Postural Stability. North American Society for the Psychology of Sport and Physical Activity (NASPSPA), Virtual Conference, Abstract published in: Journal of Sport and Exercise Psychology Vol. 42, Issue S1, 58, doi: `https://doi.org/10.1123/jsep.2020-0172`.

Symanzik, J., Li, C., Zhang, B., Studenka, B., McKinney, E., 2017. Eye-Tracking in Practice: A First Analysis of a Study on Human Postures. 2017 JSM Proceedings. American Statistical Association, Alexandria, VA, 2212–2226.

Symanzik, J., McKinney, E., Studenka, B., Bean, B., Athens, M., Hansen, M., 2018. Eye-Tracking in Practice: Results from a Study on Human Postures. 2018 JSM Proceedings. American Statistical Association, Alexandria, VA, 2696–2706.

Upton, G. J., 1992. Fisher's Exact Test. Journal of the Royal Statistical Society: Series A (Statistics in Society) 155 (3), 395–402.

Voronoi, G., 1909. Nouvelles Applications des Paramètres Continus à Théorie des Formes Quadratiques. Deuxième Mémoire. Recherches sur les Paralléloèdres Primitifs. Journal für die Reine und Angewandte Mathematik (Crelles Journal) 1909 (136), 67–182.

Wästlund, E., Shams, P., Otterbring, T., 2018. Unsold is Unseen… Or Is It? Examining the Role of Peripheral Vision in the Consumer Choice Process Using Eye-Tracking Methodology. Appetite 120, 49–56.