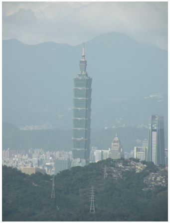
Lessons Learned from an Eye-Tracking Study on Human Postures

Jürgen Symanzik*
with
Joanna Coltrin, Chunyang Li, Eric McKinney,
Sarah Schwartz & Breanna Studenka

*Department of Mathematics and Statistics
Utah State University, Logan, UT, USA
e-mail: juergen.symanzik@usu.edu

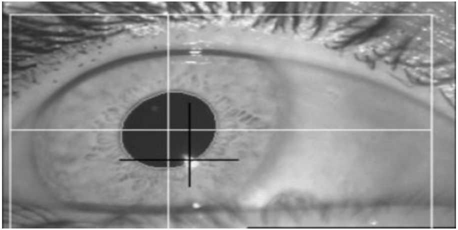


December 14, 2023

Outline

Development

- Eye-trackers first built in the late 1800s.
- Eye-tracking techniques developed rapidly during the past century.
- Video-based pupil and corneal reflection tracking method are the dominating eye-tracking method since the early 1990s.



The identified pupil (white cross-hair) and corneal reflection (black cross-hair). (Previously published in Holmqvist et al. 2011).

Applications

- **Education:** Solving problems, classroom presentations, reading, and looking at graphics.
- **Usability Research:** Reading behaviors online, searching, scanning online information, and web page design.
- **Sports:** Studying basic technical mistakes in hand-eye coordination and how to optimize performance in soccer, table tennis, shooting, hockey, and baseball.
- **Psychology:** Understanding how people gather information visually and how information is processed, e.g., research in autism.
- **Marketing:** Understanding of consumers' willingness to purchase goods.

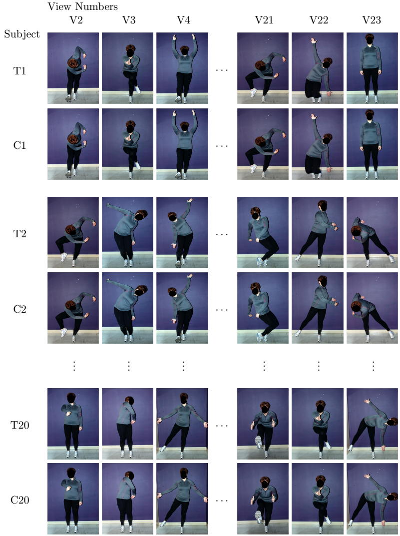
Primary Research Question

The USU Posture Study

- Does judging the action capabilities of another person depend on one's own experiences?
- Background: Action anticipation must be present when interacting with others (e.g., to avoid collisions, pass something on to someone, etc.).
- Motivated by research in the Kinesiology and Health Science Department at Utah State University (USU).
- See Symanzik et al. (2017) for details.

Experimental Setup

- 40 Subjects:
  - 20 in Treatment Group (Yoga).
  - 20 in Control Group (Non-Yoga).
- Balance measure.
- Initial calibration (View V1).
- 22 Postures (Shown in random order).
- How long do you think this person can hold this posture?
- Final calibration (View V24).
- Force plate synchronized with eye-tracker.



Anticipated Outcomes

- Those with extensive yoga experience will judge an actor to be more stable than those without stability-specific experience.
- **The visual information (i.e., viewing patterns) used to judge stability will differ between different groups of individuals with unique action experiences.**

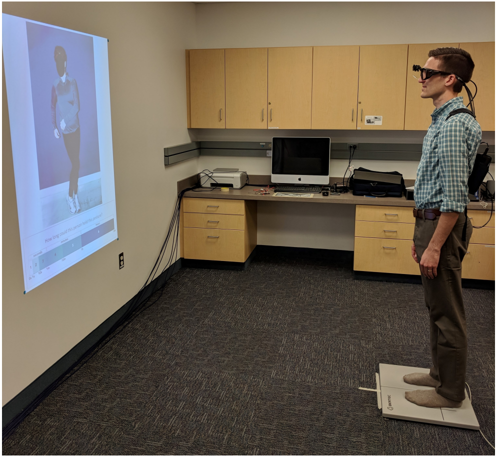
Statistical Analyses

- What are **within groups** similarities /differences (if any) of the viewing patterns for each posture / for all postures?
- What are **between groups** similarities / differences (if any) of the viewing patterns for each posture / for all postures?

- Two main analyses:
  - Large-scale analyses (see McKinney (2022) for details).
  - AOI-based analyses (see Coltrin (2022) for details).

Shooting Scene No.1

Shooting Scene No.2



Calibration of Gaze Points

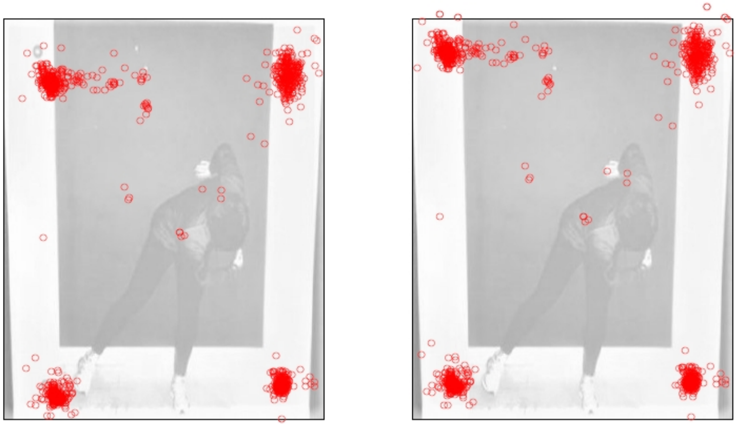


Figure: Gaze points from participant C7 for the calibration photo V1 and V24 before (left) and after (right) the affine transformation.

Calibration Contours

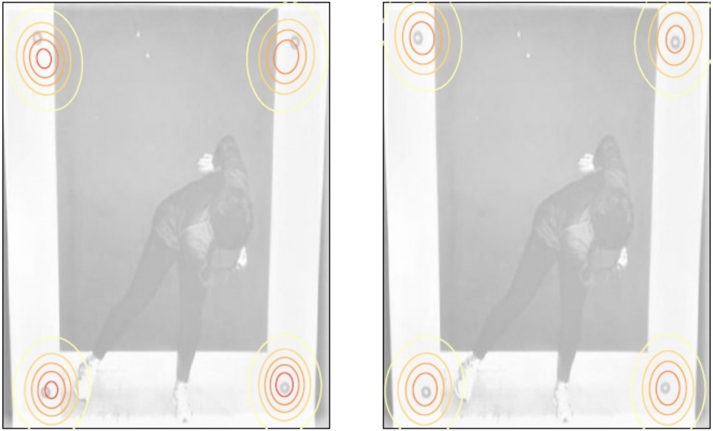


Figure: Contour plots of gaze points from all participants for the calibration photo V1 and V24 before (left) and after (right) transformation.

The Syrjala (1996) Test

- Tests for a difference between the spatial distributions of two populations.
- Sensitive to differences in the way the populations are distributed across the study area.
- Insensitive to differences in abundance between the two populations.
- Frequently used for wildlife sample surveys and epidemiology; rarely used for eye-tracking so far (e.g., Chetverikov et al., 2018).

The Modified Syrjala Test (McKinney 2022)

- Overcomes known limitations of the original Syrjala test by:
  - Using original locations instead of binning.
  - Making rotational modifications.
  - Making toroidal shift modifications.
  - Making adjustments of weights for different numbers of gaze points collected for each subject in the Control and Treatment Groups.
  - Setting thresholds for the number of toroidal shifts for speedup.

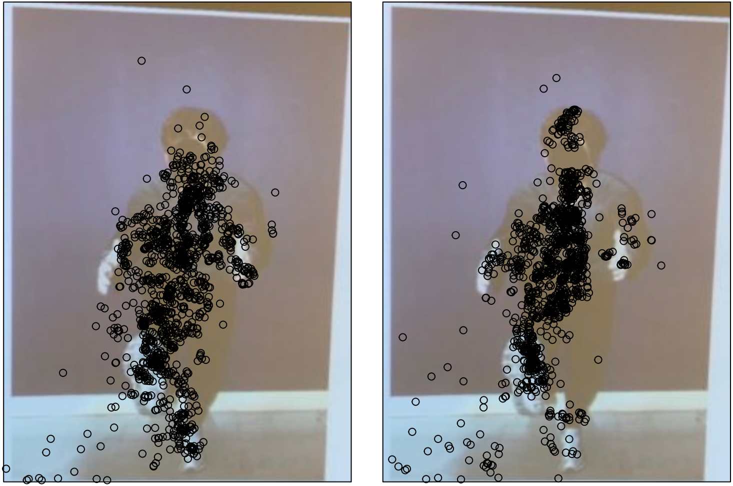


Figure: Aggregated gaze point scatterplots for the Treatment (left) and Control (right) Groups for posture ID 2.

Table: USU Posture Study aggregated group test results. All p-values were identical except for * where the threshold test resulted in 0.02.

Posture ID	n_1	n_2	p-values	Comp. Time (in hours) of Proportional Test	Comp. Time (in hours) of Threshold Test
1	1091	1297	0.01	2.10	0.31
2	1289	1281	0.01	2.49	0.37
3	1356	1346	0.01	2.87	0.39
4	1452	1479	0.01	3.46	0.42
5	1900	1523	0.01	5.00	0.52
6	1716	1360	0.01*	3.89	0.45
7	1359	1246	0.01	2.63	0.34
8	1089	1186	0.01	1.87	0.32
9	1313	1242	0.01	2.54	0.35
10	1624	1215	0.01	3.26	0.41
11	1184	1443	0.01	2.66	0.35
12	1436	1337	0.01	3.64	0.38
13	1869	1768	0.01	6.45	0.63
14	1421	1202	0.01	3.08	0.48
15	1396	1343	0.01	3.64	0.46
16	1541	1352	0.01	3.97	0.42
17	1661	1508	0.01	4.80	0.59
18	1746	1367	0.01	4.88	0.46
19	1126	1108	0.01	2.34	0.32
20	1348	1374	0.01	3.53	0.37
21	1476	1597	0.01	4.33	0.44
22	1501	1517	0.01	4.04	0.43

Table: USU Posture Study pairwise within group test result examples. Two significant and two non-significant tests within each group were randomly selected as examples.

Group	Subj. 1 ID	Subj. 2 ID	Post. ID	n_1	n_2	p-values	Comp. Time (in secs)
Treatment	8	18	20	56	53	0.01	0.67
Control	3	4	22	61	44	0.01	0.67
Treatment	9	17	11	80	30	0.10	0.68
Control	8	12	19	21	21	0.08	0.10

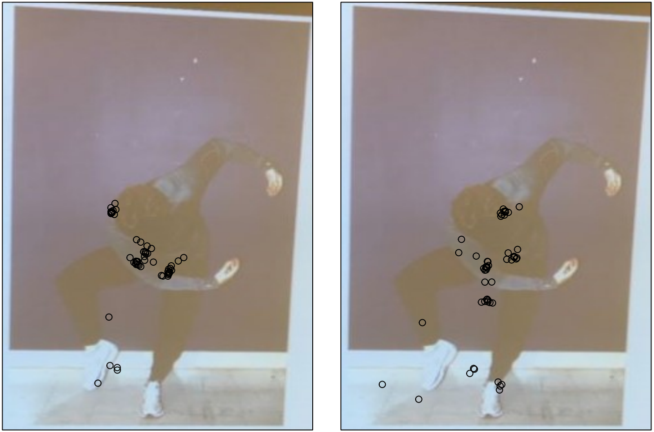


Figure: Scatterplots of the gaze points for subject ID 8 (left) and subject ID 18 (right) within the Treatment Group for posture ID 20. The test result was significant (p-value = 0.01) for this comparison.

Figure: Scatterplots of the gaze points for subject ID 8 (left) and subject ID 12 (right) within the Control Group for posture ID 19. The test result was non-significant (p-value = 0.08) for this comparison.

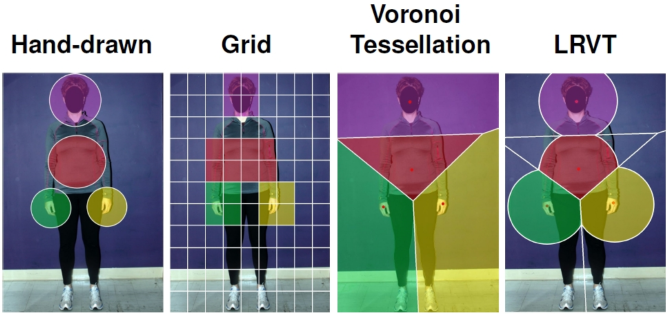
Summary of Large-Scale Analyses

- All 22 group-wise tests were statistically significant.
- Out of the 4,180 individual tests within each group, only 35 were non-significant within the Treatment Group (99.16% significant), and only 54 were non-significant in the Control Group (98.71% significant).
- Furthermore, the largest non-significant p-values within the Treatment and Control Groups are only 0.25 and 0.37, respectively, and many of the remaining non-significant p-values tend to be close to the significance level of 0.05.
- Omission of a multiple testing correction due to richness of significant tests.
- The subjects within each respective group exhibit mostly heterogeneous gaze point patterns.

Areas of Interest

According to Hessels et al. (2016):

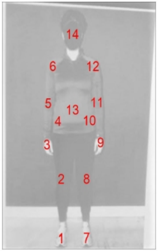
- Areas of interest (AOIs) are predefined areas of an image.
- AOIs can be hand-drawn or defined systematically.
- Four main types of AOIs: Hand-drawn, Grid, Voronoi Tessellation (VT), and Limited-Radius Voronoi Tessellation (LRVT).



AOI Center Definitions

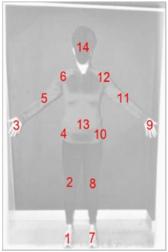
- AOI centers were defined by a kinesiology expert.
- AOI centers include major joints, the torso, and the head.

Posture ID 1



Posture ID 5

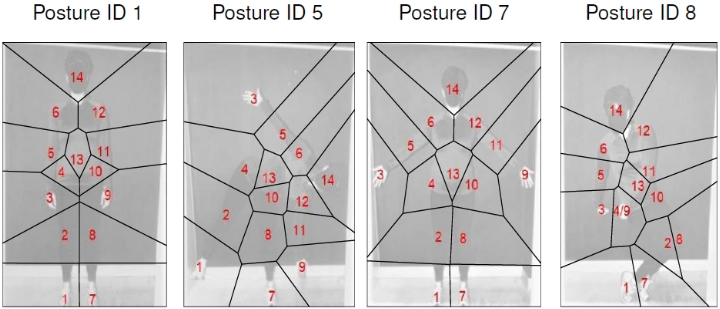
Posture ID 7



Posture ID 8

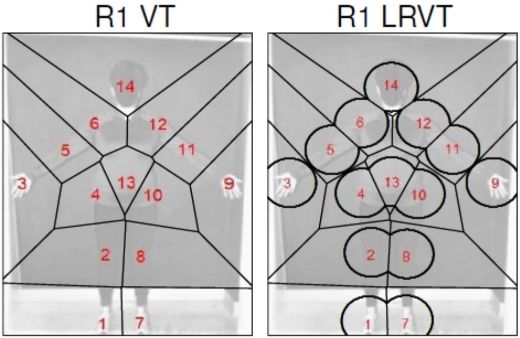
Defining VT AOIs

- Voronoi Tessellations were applied to the postures as depicted below.



Defining LRVT AOIs

- Limited-Radius Voronoi Tessellations were applied to the postures as depicted below for Posture ID 7.

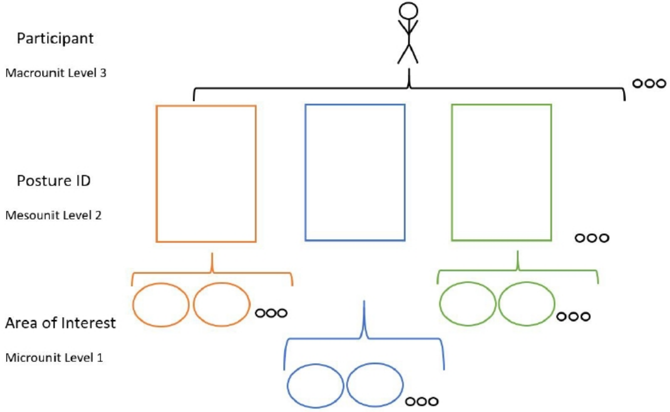


Methods for AOI Analyses

Multilevel Analyses via Generalized Linear Mixed Effects Models (GLMMs) using:

- Wald Chi-Square Test (to determine if the model fit predicts the gaze point count better with fixed effects than without).
- Poisson Regression.
- Negative Binomial Regression.
- Zero-Inflated Poisson (ZIP) Regression.
- Zero-Inflated Negative Binomial (ZINB) Regression.

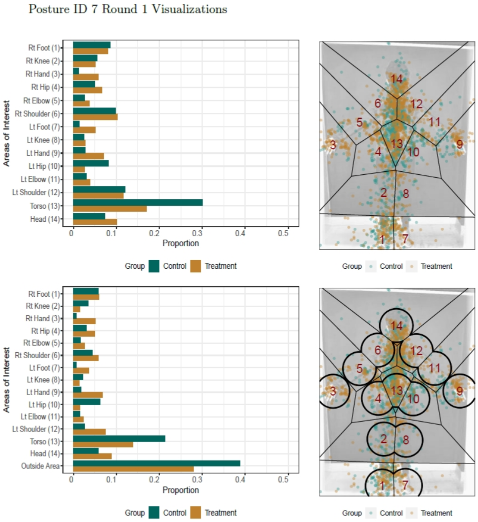
Multilevel Structure



Hypotheses

- Aggregated analyses:
  - H_0: There is no association between Group and AOI.
  - H_A: There is an association between Group and AOI.
  - These tests are analyzed for each Posture ID separately.
- Multilevel analysis:
  - H_0: The fixed effects of Group, AOI, and the cross-level interaction between Group and AOI all are equal to 0, i.e., they do not significantly improve the ability to predict the counts.
  - H_A: At least one of these parameters is different from 0, i.e., at least one of them is significant when predicting the gaze point counts.

Visual Results for VT-based and LRVT-based AOIs for ID 7



Test Results for VT-based and LRVT-based AOIs for ID 7

- ID 7 shows some of the most significant p-values.
- Graphically, for VT-based AOIs, main differences can be seen for Rt Hand (3), Lt Hand (9), and Torso (13).
- For LRVT-based AOIs, nearly 40% (for Control Group) and nearly 30% (for Treatment Group) of all gaze points fall into the “Outside Area”, i.e., outside of the AOIs.
- Overall, each test shows significant differences in Treatment vs. Control proportions within AOIs for all 22 postures.

Visual Pairwise Predictions for all VT-based AOIs

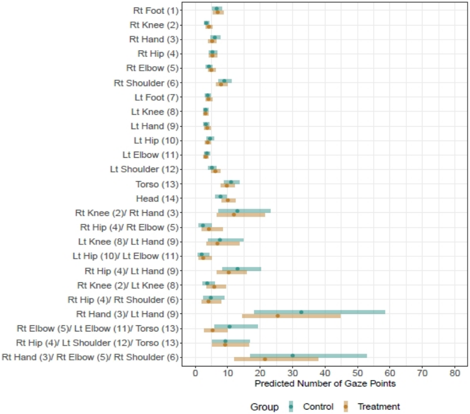


Figure: The pairwise comparisons of the predicted counts by AOI and Group visualize the insignificant findings from the Wald Test for the ZINB Regression model.

Visual Pairwise Predictions for all LRVT-based AOIs

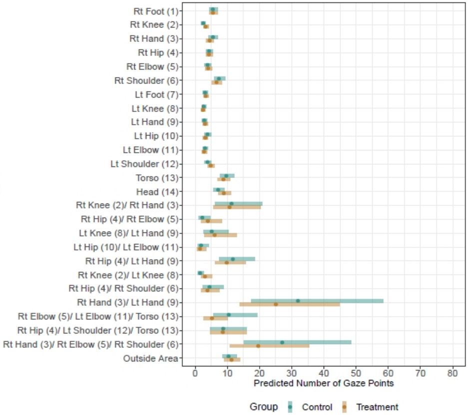


Figure: The pairwise comparisons of the predicted counts by AOI and Group visualize the insignificant findings from the Wald Test for the ZINB Regression model. Note the “Outside Area”.

Conclusions (1)

- The aggregated analyses give a high level of statistical significance that there is an association between Group and AOI, refining the results from the large-scale analyses.
- The GLMMs give conflicting results about the Group $\times$ AOI interaction.
- The ZINB without interaction model (best fit) suggests that while there may not be significant differences between Groups, there are highly statistically significant gaze-point differences between AOIs.

Conclusions (2)

- Visually, it appears that participants spent more time looking at the head and torso as compared to the extremities.
- Different results for Poisson and ZIP models (Negative Binomial did not converge).
- Different results for VT-based and LRVT-based AOIs.
- Overall, results dependent on two factors: how the AOIs are defined and which regression model is used.

Discussion

- Considerable amount of noise in the eye-tracking-based gaze point locations may mask at which AOI a viewer looked in case of nearby AOIs.
- Alternative analyses based on fewer, but better separated AOIs that were created based on the joint gaze point distributions yield different results.
- LRVT-based AOIs may help to further filter out noise from gaze points.
- Suggestion: Start with fewer AOIs immediately that are better separated.

References

- Chetverikov, A., et al. (2018). Implicit Processing during Change Blindness Revealed with Mouse-Contingent and Gaze-Contingent Displays. *Attention, Perception, & Psychophysics*, 80(4):844–859, `https://doi.org/10.3758/s13414-017-1468-5`.
- Coltrin, J. D. (2022). *Defining Areas of Interest Using Voronoi and Modified Voronoi Tesselations to Analyze Eye-Tracking Data*, MS Thesis, Utah State University, All Graduate Theses and Dissertations, 8508, `https://doi.org/10.26076/b042-bcdb`.
- Hessels, R. S., et al. (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, `https://doi.org/10.3758/s13428-015-0676-y`.
- Holmqvist, K., et al. (2011). *Eye Tracking: A Comprehensive Guide to Methods and Measures*, Oxford, New York: Oxford University Press.
- McKinney, E. D. (2022). *Extensions to the Syrjala Test with Eye-Tracking Data Analysis Applications in R*, PhD Thesis, Utah State University, All Graduate Theses and Dissertations, 8539, `https://doi.org/10.26076/ea47-6085`.
- Symanzik, J., et al. (2017). Eye-tracking in Practice: A First Analysis of a Study on Human Postures. In *2017 JSM Proceedings*, Alexandria, VA: American Statistical Association. (CD).

- **Questions ?!? —**

- or e-mail: `juergen.symanzik@usu.edu`