UNIVERSITY OF TAMPERE

Unit testing fixtures for development of automatic gaze analysis methods

Method description

University of Tampere
School of Medicine
Infant Cognition Laboratory

Aug 17th 2015
Updated Dec 31th 2015

Project team
Akseli Palén <akseli.palen@staff.uta.fi> – Producer and study participant
Jussi Kaatiala <jussi.kaatiala@staff.uta.fi> – Analysis software and setup designer
Mika Tammi <mikatammi@gmail.com> – Camera operator and laboratory assistant
Jukka Leppänen <jukka.leppanen@staff.uta.fi> – Manager

Table of Contents

1. Introduction

In the development of automatic analysis methods for infant gaze studies, we design and build complex algorithms and software tools. To explore, measure and ensure the function of the tools, such as their accuracy, precision, and robustness, we have to test them against quality data that we fully understand already.

This document describes such data for gaze saccades and fixations and how it was produced. We go through the laboratory environment, devices, their positioning, and calibration. We also provide visualizations of the results. However, the resulting recordings are not provided in this document but as separate files.

2. Devices

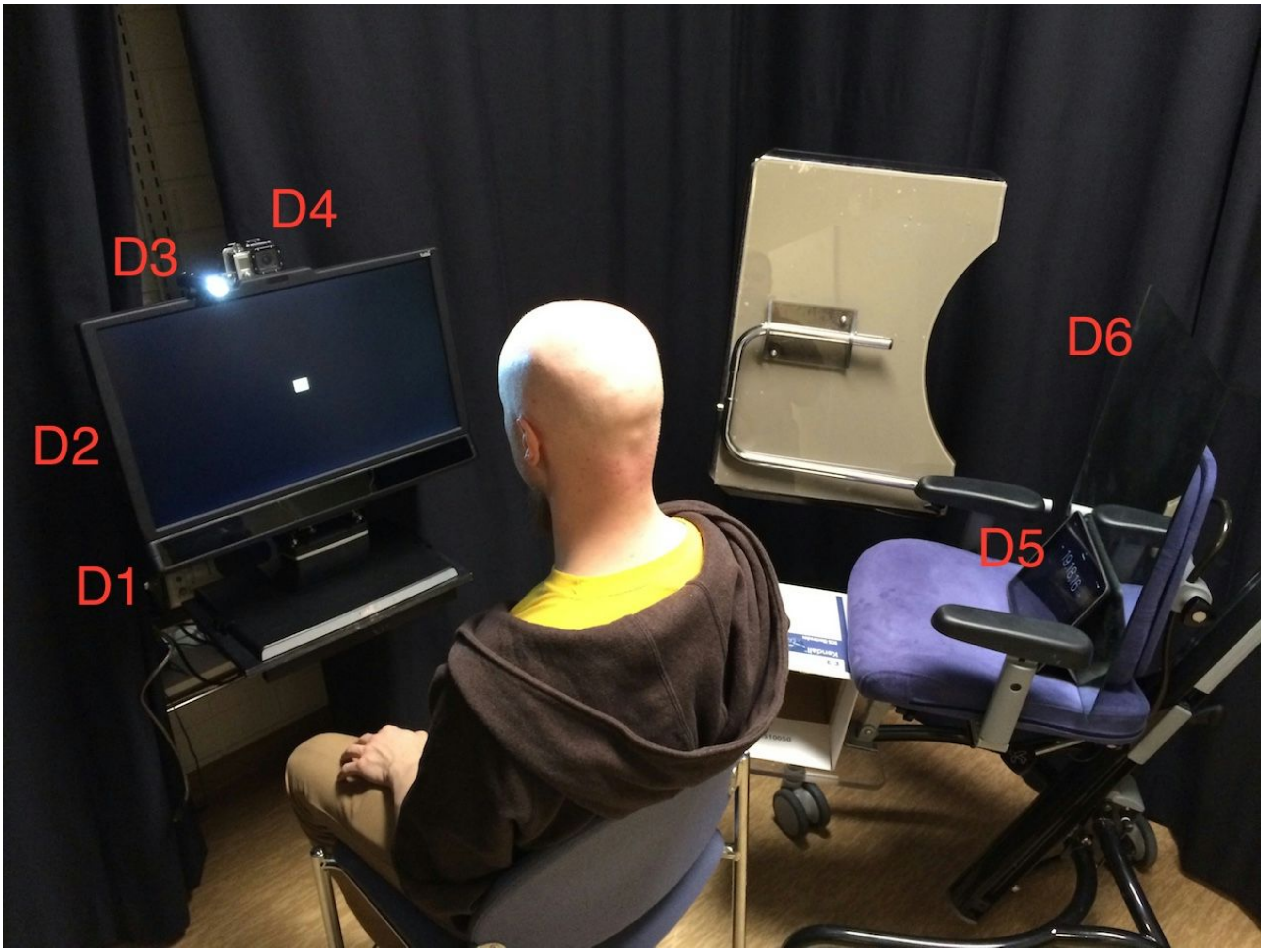


Figure DEVICES. The laboratory setup and devices. The labeled devices are described in Table DEVICES.

Table DEVICES. The devices that were used to produce the data.

ID	Device	Purpose
D1	Tobii eye tracker	Captures the gaze position on screen
D2	Tobii display	Presents the visual stimuli
D3	Led Lenser LED flashlight	Brightens the eyes to improve contrast in video recording
D4	GoPro video camera	Captures the eye movements, clock, and what happens on the screen.
D5	iPad stopwatch app	Gives reference time that enables us to detect possible drift in the camera framerate.
D6	Mirror	Reflects the events on the screen for the video camera.

3. Positioning

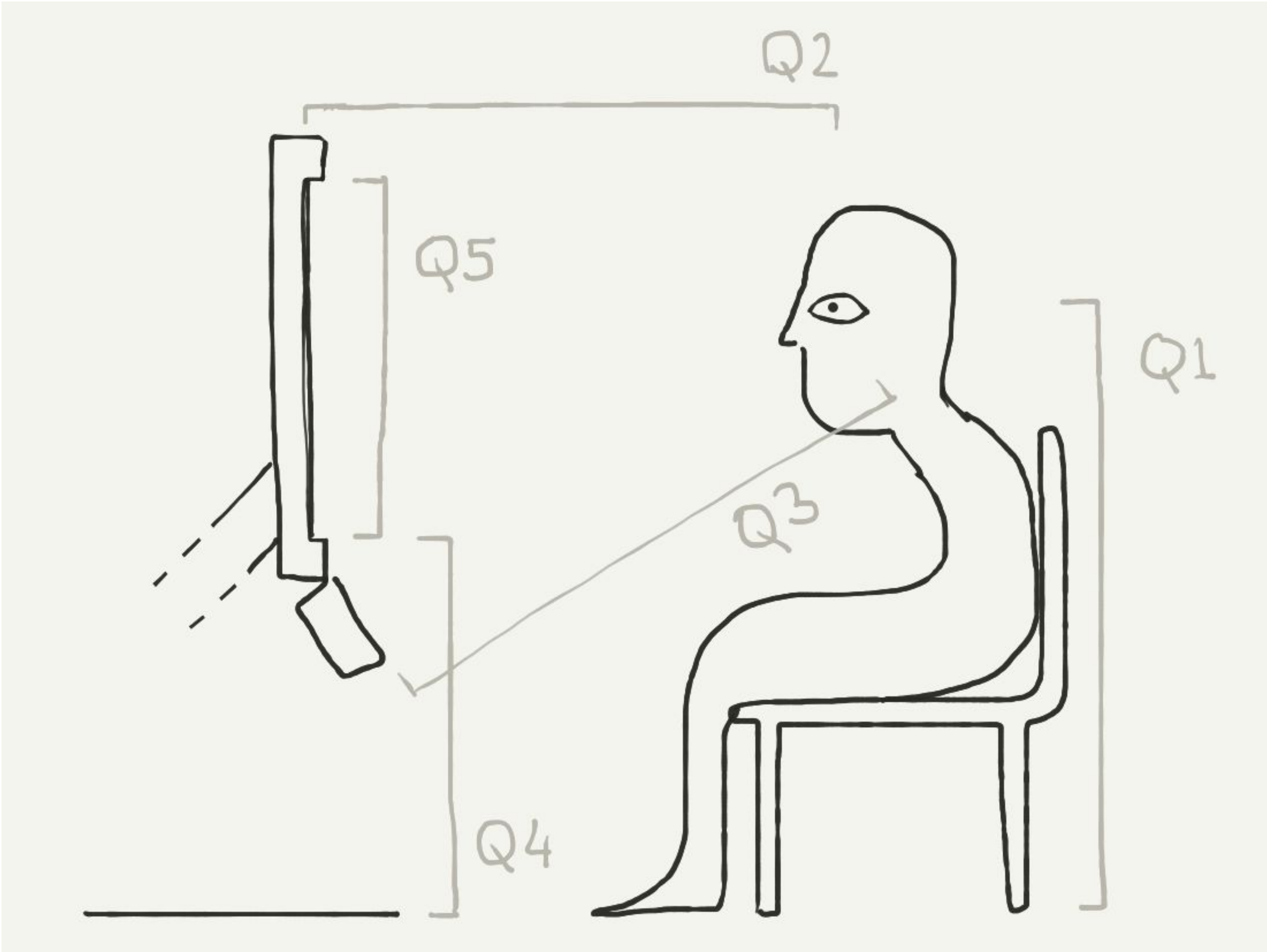


Figure POSITIONING. Quantities of the physical setup. The labeled quantities are described in Table POSITIONING.

Table POSITIONING. Quantities and measurements that describe the physical setup. The numbers were collected manually by using standard tape measure.

ID	Quantity description	Min (mm)	Max (mm)	Avg (mm)
Q1	Pupil height from the floor	1170	1200	1185
Q2	Pupil distance to the screen (orthogonal)	610	630	620
Q3	Pupil distance to the eye tracker (orthogonal)	660	670	665
Q4	Height of the bottom edge of the screen panel from the floor.	884	886	885
Q5	Visible screen panel height	287	289	288
Q6	Visible screen panel width	509	511	510

4. Animation

To create the data we asked the participant to stare at the middle of a white rectangle on black background (Figure STIMULUS) and follow it if it chages place on screen.

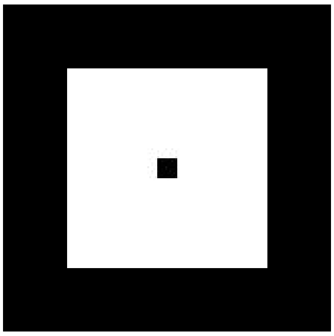


Figure STIMULUS. The white rectangle with a black dot. The dot helps to determine the middle point.

5. Video recording

Figure FRAME. A single frame from the recording. The participant is looking at the white square at the upper right corner of the screen. The square and the screen are visible to the camera through the mirror. The stopwatch timer can be seen below the mirror, displaying 4 minutes 56.58 seconds.

Video camera

- Model: GoPro HERO3 Silver
- Framerate: 120 Hz
- Resolution: 848x480 pixels

Video file

QuickTime Player 10.4 Movie Inspector tells us the following:

- Format: GoPro AVC encoder, 848x480 AAC, 48000 Hz, Stereo (L R)
- FPS: 119.93
- Data Size: 180.5 MB

Framerate

It is possible that the advertized framerate of the camera is not exactly 120 Hz, what could produce systematic, linearly increasing error when events are timed. To rule this out, we compare the frame indices and stopwatch times. We use QuickTime Player 7 to inspect the video frame by frame.

Frame indices start from 0. At frame 4, the stopwatch has just changed from 3m21.97s to 3m22.02s. At frame 11777 the clock has just changed from 5m0.07s to 5m0.11s. Therefore 11773 frames were elapsed during the confidence interval of [98.05, 98.14] seconds. This gives us measured framerate of [119.96, 120.07] Hz. Counting the time from the frames alone with 120.00 Hz, yields 98.11s which is between the confidence interval. Counting the frames from the stopwatch times alone with 120.00 Hz, yields [11766, 11777] frames to which the 11773 frames well fit. Therefore, **we can safely use frame indices** to time recorded events.

6. Raw events from video recording

We enhanced the the interframe differences in the recorded video to better determine the frame indices of events. We also cropped the video for quicker computation time. Figure DIFF gives an example of the resulting video.

Figure DIFF. The differences in pixel values between video frames are emphasized by the magenta color. As the eyes move, especially the edges of the irises become colored. The differences on the wall are caused by the rapidly flashing fluorecent lamp. The face is illuminated with LED light that does not produce such flashing effect.

From the video we extracted events such as saccade beginnings and stimulus movements. The events and their frame indices are listed in Table EVENTS.

Table EVENTS. The events are in chronological order. LCB = lower confidence bound, UCB = upper confidence bound. Saccades back to the middle are omitted.

Event	Frame index LCB	Frame index UCB
Square appears in the middle	1312	1312
Eyelids closed	1389	1392
Eyelids half-open	1400	1402
Eyelids closed	1411	1440
Eyes fixated to the square		
Eyelids closed	2009	2016
Eyelids closed	2117	2128
Eyelids closed	2283	2297

Eyelids closed	2423	2435
Square disappears from the middle	2514	2516
Square appears in the middle	2784	2785
Square moves to lower-right corner	3024	3025
Saccade starts	3075	3075
Saccade ends	3079	3079
Square moves to the middle	3506	3507
Square moves to the lower-left corner	3748	3749
Saccade starts	3777	3777
Saccade ends	3783	3783
Square moves to the middle	4230	4231
Square moves to the upper-left corner	4472	4473
Saccade starts	4490	4490
Saccade ends	4496	4497
Eyelids closed	4704	4706
Square moves to the middle	4952	4952
Eyelids closed	4987	4992
Eyelids closed	5060	5072
Square moves to upper-right corner	5194	5194
Saccade starts	5216	5216
Saccade ends	5223	5223
Square moves to the middle	5676	5676
Square moves to the lower-right corner	5919	5919
Saccade starts	5948	5948
Saccade ends	5953	5954
Square moves to the middle	6401	6401
Square moves to the lower-left corner	6643	6643
Saccade starts	6658	6659
Saccade pause	6662	6663

Saccade unpause	6717	6718
Saccade end	6721	6721
Square moves to the middle	7125	7125
Square moves to the upper-left corner	7366	7367
Saccade starts	7390	7391
Saccade pause	7398	7399
Saccade unpause	7459	7460
Saccade end	7463	7464
Eyelids close	7604	7609
Eyelids close	7663	7670
Square moves to the middle	7848	7849
Square moves to the upper-right corner	8090	8091
Saccade starts	8113	8113
Saccade pause	8119	8120
Saccade unpause	8179	8180
Saccade end	8183	8184
Square moves to the middle	8570	8571
Square moves to the lower-right corner	8813	8814
Saccade starts	8836	8837
Multiple small saccades		
Saccade ends	8908	8910
Square moves to the middle	9295	9296
Square moves to the lower-left corner	9538	9538
Saccade starts	9558	9559
Multiple small saccades		
Saccade ends (hard to determine)	9599	9600
Square moves to the middle	10019	10020
Square moves to the upper-left corner	10261	10261
Saccade starts	10290	10291

Multiple small saccades		
Saccade ends (hard to determine)	10357	10359
Eyelids close	10473	10480
Square moves to the middle	10743	10743
Square moves to the upper-right corner	10985	10985
Saccade starts	11016	11016
Saccade ends	11076	11080
Square disappears (hard to determine)	11467	11470

7. Saccade parameters from the raw video events

Table. Calculated millisecond times are based on the framerate of 120.0 Hz. F# = Frame index, Fs = number of frames, SRT = Saccadic Reaction Time, SD = saccade duration.

Trial #	Stimulus F#	Sacc. start F#	Sacc. end F#	SRT Fs	SD Fs	SRT (ms)	SD (ms)
0	3024.5	3075.0	3079.0	50.5	4.0	420.8	33.3
1	3748.5	3777.0	3783.0	28.5	6.0	237.5	50.0
2	4472.0	4490.0	4496.5	18.0	6.5	150.0	54.2
3	5194.0	5216.0	5223.0	22.0	7.0	183.3	58.3
4	5919.0	5948.0	5953.5	29.0	5.5	241.7	45.8
5	6643.0	6658.5	6721.0	15.5	62.5	129.2	520.8
6	7366.5	7390.5	7463.5	24.0	73.0	200.0	608.3
7	8090.5	8113.0	8183.5	22.5	70.5	187.5	587.5
8	8813.5	8836.5	8909.0	23.0	72.5	191.7	604.2
9	9538.0	9558.5	9599.5	20.5	41.0	170.8	341.7
10	10261.0	10290.5	10358.0	29.5	67.5	245.8	562.5
11	10985.0	11016.0	11078.0	31.0	62.0	258.3	516.7

8. Stimulus durations from the raw video events

Table. Note the reading precision of ±0.5 frames ≈ ±4.2 ms. Abbreviations: Fs = number of frames, "=>" = implied from the other values.

Trial #	Foreperiod Fs	Target Fs	=> Fore ms	=> Target ms
0	240.0	482.0	2000.0	4016.7
1	242.0	482.0	2016.7	4016.7
2	241.5	480.0	2012.5	4000.0
3	242.0	482.0	2016.7	4016.7
4	243.0	482.0	2025.0	4016.7
5	242.0	482.0	2016.7	4016.7
6	241.5	482.0	2012.5	4016.7
7	242.0	480.0	2016.7	4000.0
8	243.0	482.0	2025.0	4016.7
9	242.5	481.5	2020.8	4012.5
10	241.5	482.0	2012.5	4016.7
11	242.0	483.5	2016.7	4029.2 (outlier)

9. Stimulus durations from eye tracker and comparison

We compare how much the timestamps from the tracker differ from the times captured on video. This allows us to ensure the values are well synchronized.

Table TRACKPERIODTBT. Foreperiod and target period durations taken **from TBT** and how much they differ from the durations in video. Units in milliseconds.

Trial #	Foreperiod	Target period	Video fore	Video target	Diff to video foreperiod	Diff to video target period
0	1996.4	4012.8	2000.0	4016.7	–3.6	–3.9
1	2016.4	4012.8	2016.7	4016.7	–0.3	–3.9
2	2013.1	3996.2	2012.5	4000.0	+0.6	–3.8
3	2013.1	4016.1	2016.7	4016.7	–3.6	–0.6
4	2013.0	4012.8	2025.0	4016.7	–12.0	–3.9
5	2013.1	4016.2	2016.7	4016.7	–3.6	–0.5
6	2013.1	4012.9	2012.5	4016.7	+0.6	–3.8
7	2013.0	3999.4	2016.7	4000.0	–3.7	–0.6
8	2013.1	4012.9	2025.0	4016.7	–11.9	–3.8
9	2013.1	4016.1	2020.8	4012.5	–7.7	+3.6
10	2013.1	4012.8	2012.5	4016.7	+0.6	–3.9
11	2013.0	3996.2	2016.7	4029.2 (outlier)	–3.7	–33.0 (outlier)

Table PERIODTBTDIFF. Aggregates without outliers. Unit is milliseconds.

Min	1996.4	3996.2		–12.0	–3.9
Max	2016.4	4016.2		+0.6	+3.6
Mean	2012.0	4009.8		–4.0	–2.3
SD	4.99	7.71		4.46	2.48

Table TRACKPERIOD. Foreperiod and target period durations taken from **raw eye-tracker data** and how much they differ from the durations in video. Units in milliseconds.

Trial #	Foreperiod	Target period	Video fore	Video target	Diff to video foreperiod	Diff to video target period
0	2000.225	4017.184	2000.0	4016.7	+0.225	+0.484
1	2016.977	4019.199	2016.7	4016.7	+0.277	+2.499
2	2016.939	4000.542	2012.5	4000.0	+4.439	+0.542
3	2016.953	4017.171	2016.7	4016.7	+0.253	+0.471
4	2016.971	4017.225	2025.0	4016.7	–8.029	+0.525
5	2017.210	4016.878	2016.7	4016.7	+0.510	+0.178
6	2016.992	4017.129	2012.5	4016.7	+4.492	+0.429
7	2017.345	4000.193	2016.7	4000.0	+0.645	+0.193
8	2016.943	4017.287	2025.0	4016.7	–8.057	+0.587
9	2016.907	4017.142	2020.8	4012.5	–3.893	+4.642
10	2016.974	4017.238	2012.5	4016.7	+4.474	+0.538
11	2016.940	4000.608	2016.7	4029.2 (outlier)	+0.240	–28.592 (outlier)

Table PERIODDIFF. Aggregates without outliers. Unit is milliseconds.

Min	2000.2	4000.2	2000.0	4000.0	–8.1	+0.2
Max	2017.3	4019.2	2025.0	4029.2	+4.5	+4.6
Mean	2015.6	4013.1	2016.0	4014.6	–0.4	+1.0
SD	4.85	7.68	6.61	6.69	4.30	1.36

As can be seen from Table PERIODDIFF, in average, foreperiod and target period last a little longer on the screen than what can be understood from the eye tracker data. Although, the difference is only half of the frame interval (8.3 ms) and therefore we can conclude that the screen and the tracker data are well synchronized.

10. Fixations

We recorded 10 seconds of fixation, represented in the 3 figures below. The slice #600-#1600 includes no invalid points and the slice #1600-#2600 includes invalid points due to blinks.

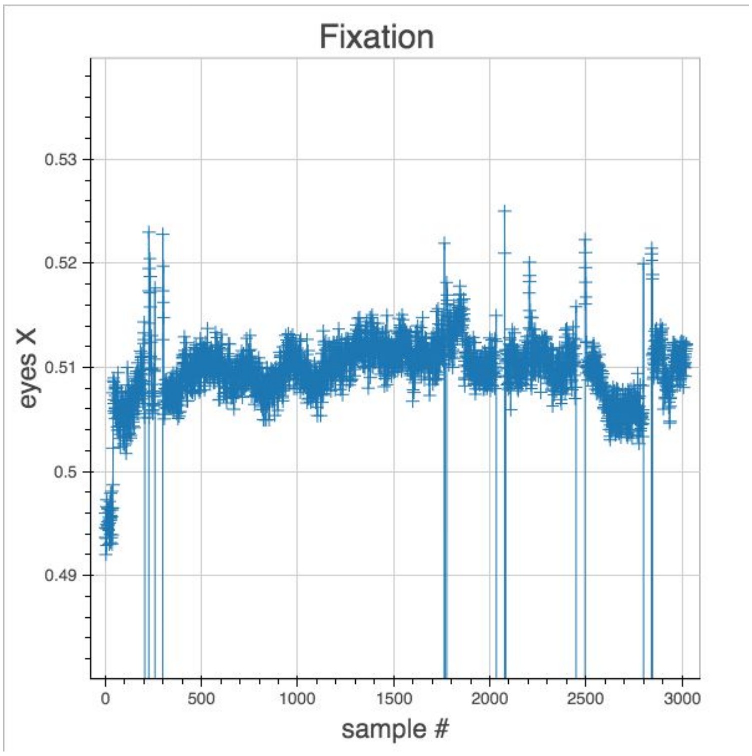


Figure. Horizontal gaze position during 10 secs @ 300 Hz

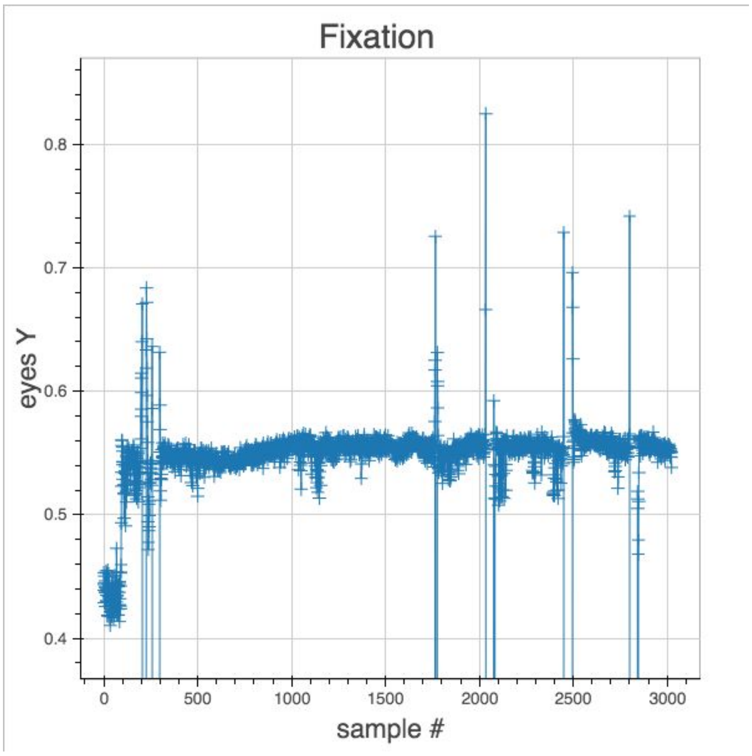


Figure. Vertical gaze position during 10 secs @ 300 Hz

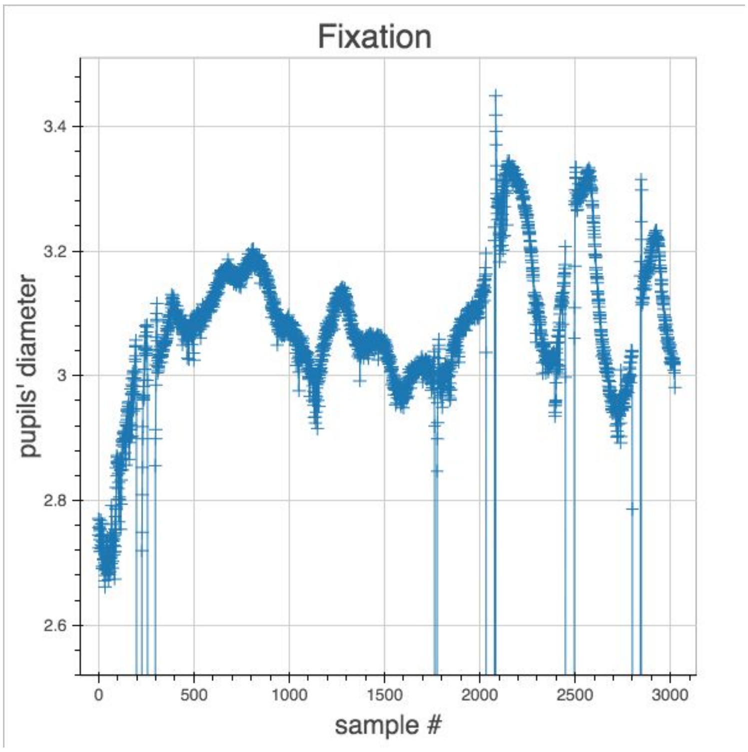


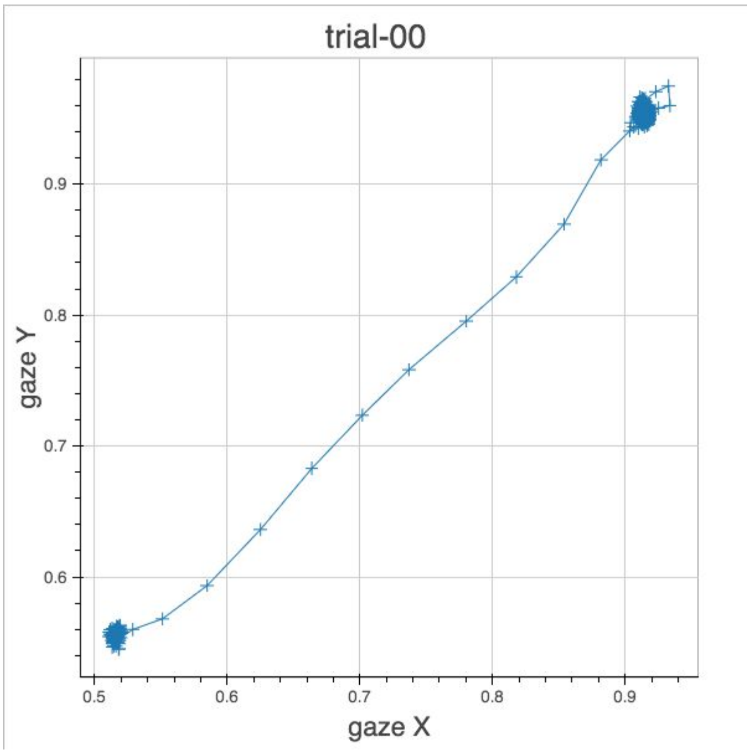
Figure. Combined pupil diameter during 10 secs @ 300 Hz

11. Saccades

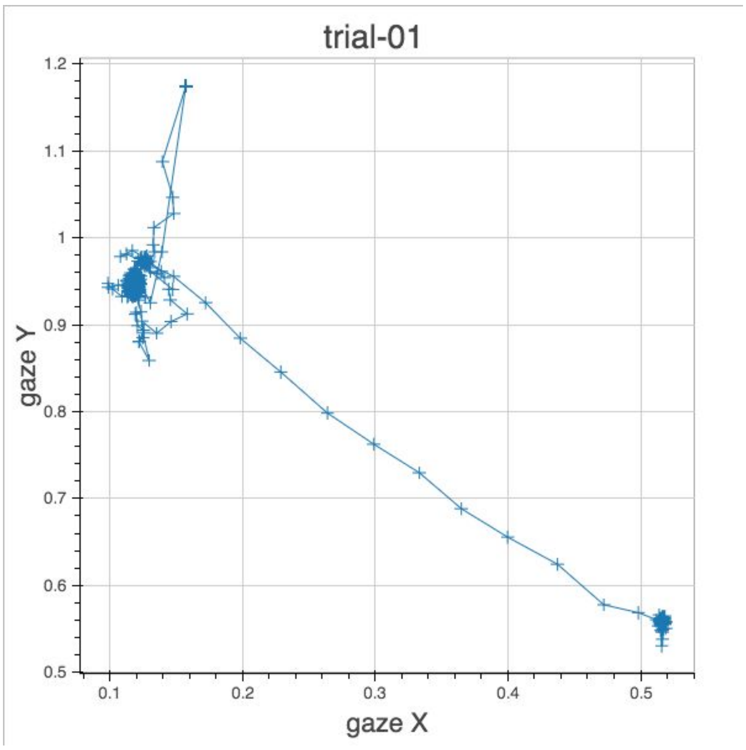
We recorded 12 saccades, consisting of the following types:

- Trials 00-03: direct saccades
- Trials 04-07: multipart saccades
- Trials 08-11: curved multipart saccades

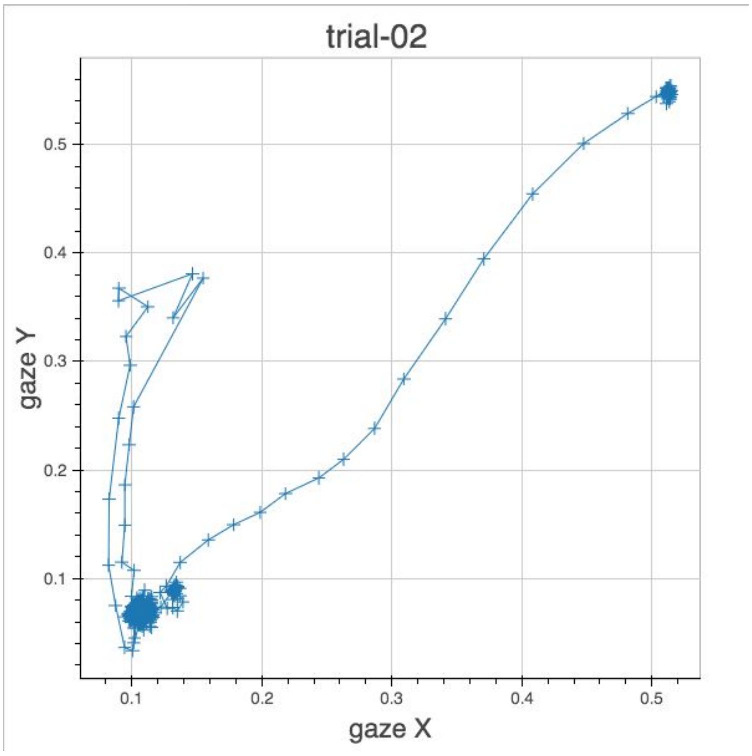
Trial 0



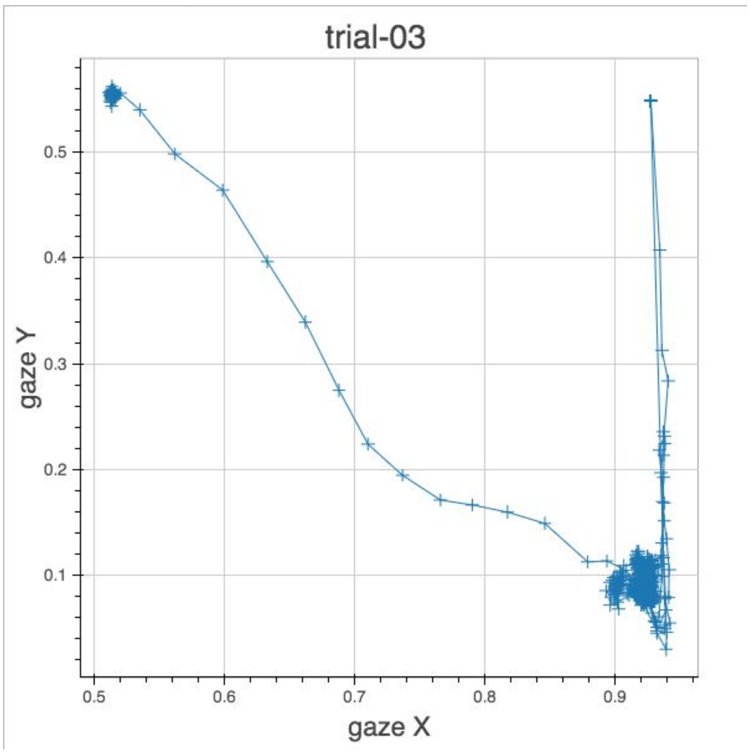
Trial 1



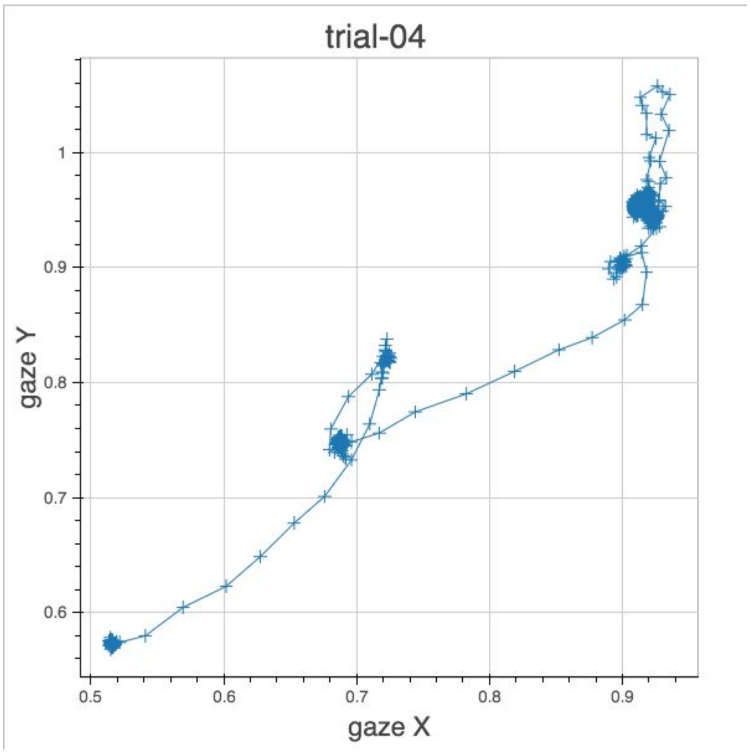
Trial 2



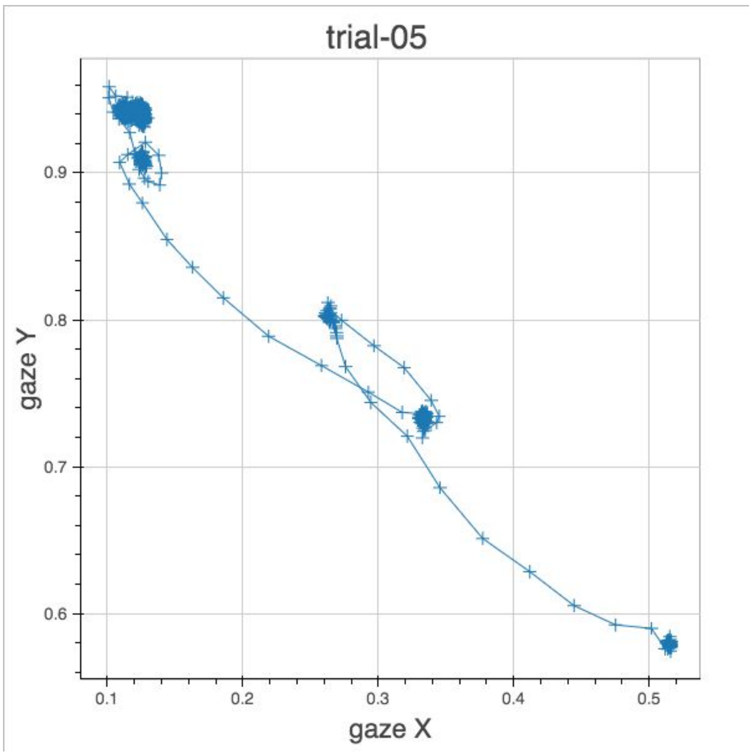
Trial 3



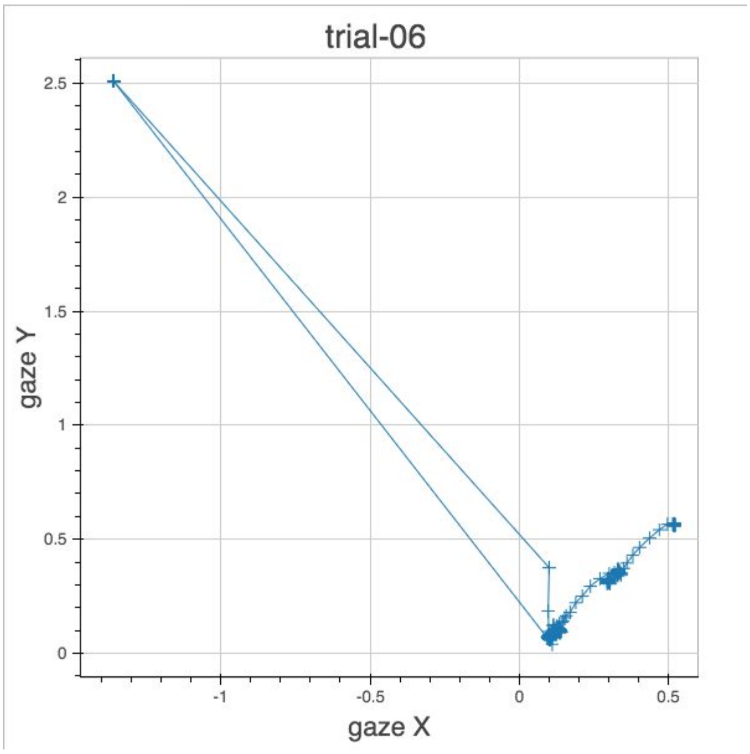
Trial 4



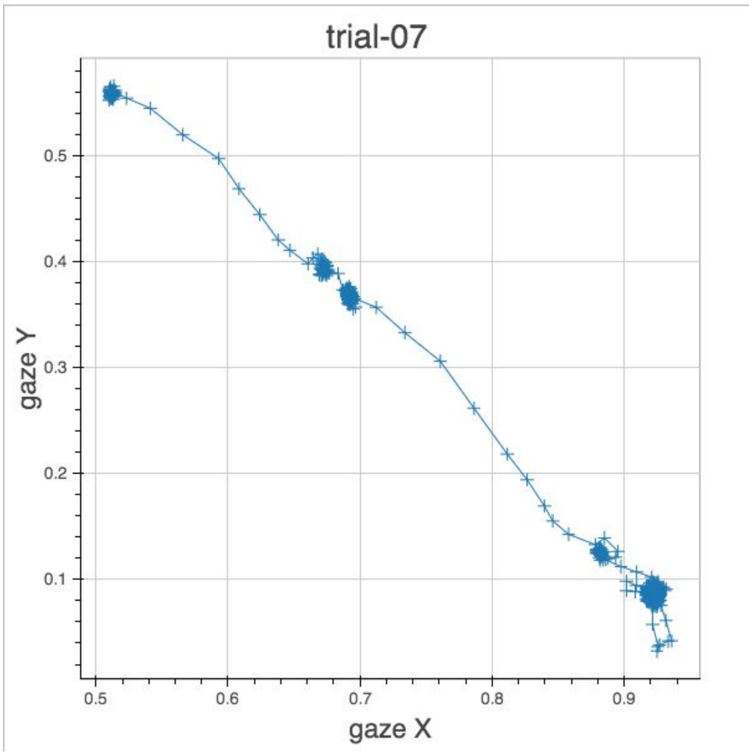
Trial 5



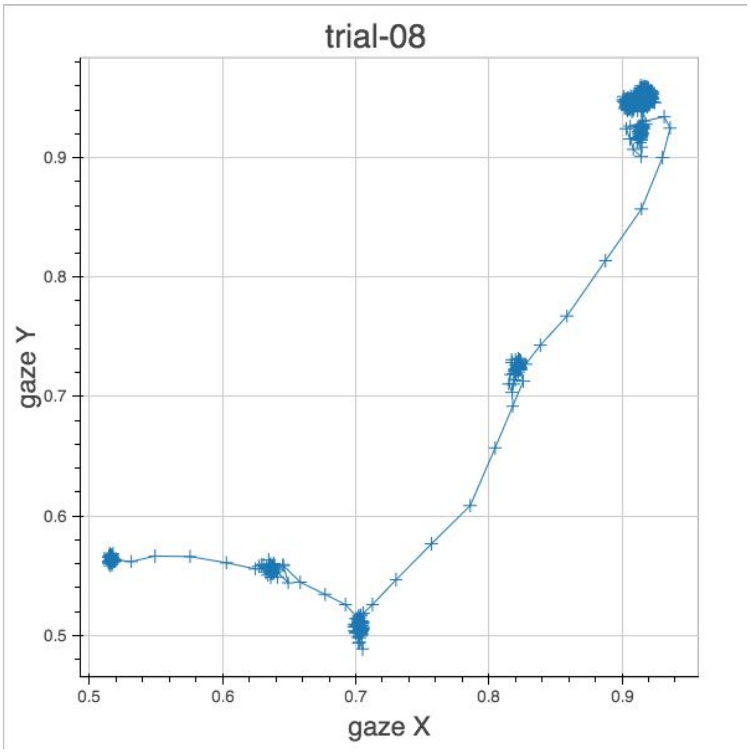
Trial 6



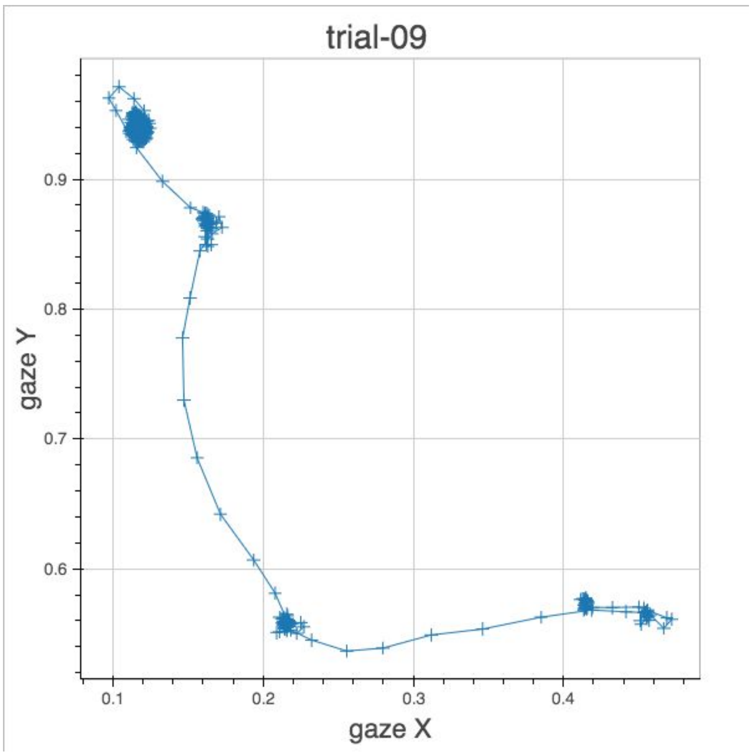
Trial 7



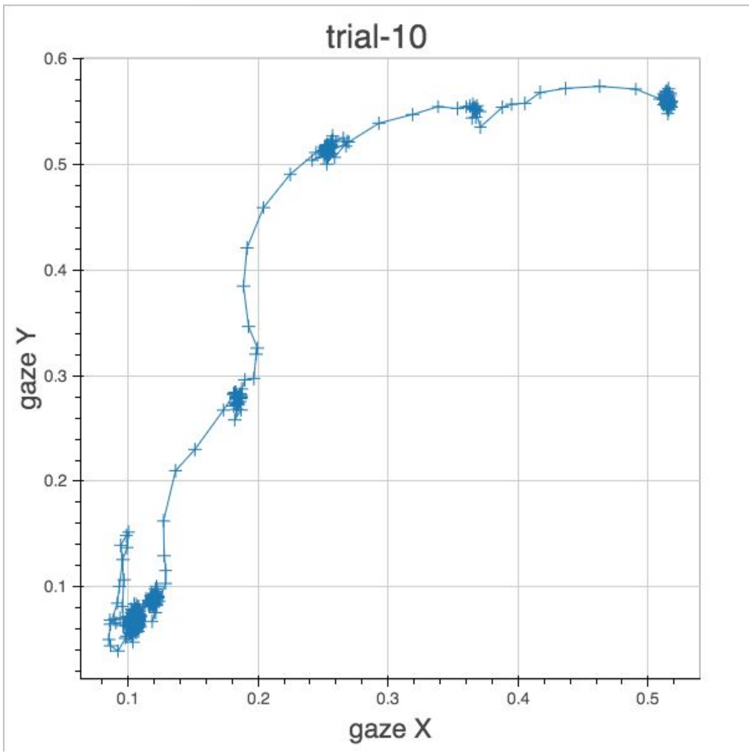
Trial 8



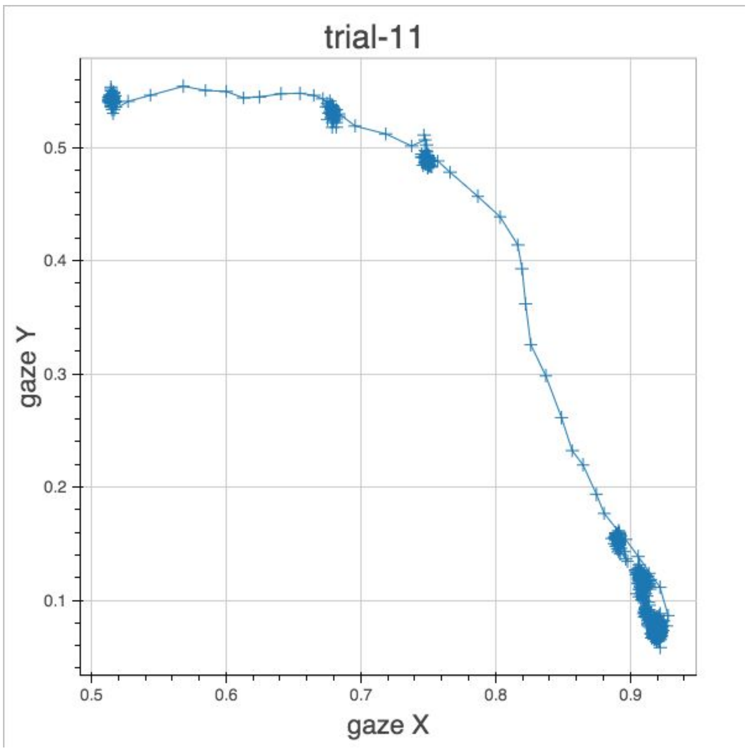
Trial 9



Trial 10



Trial 11



12. Discussion

The results provide reference data for algorithm development on the precision level of about ±10 ms. If more precise numbers were needed, we should use a camera with higher framerate. A framerate double the framerate of the eye tracker would be ideal.