Eye Trackers Comparison — Statistical Report

February 4, 2026

Contents

Variable	n	Mean [95% CI]	SD	Median (IQR)
Unrelevant / Loss (%)	67	49.9 [45.5, 54.3]	18.1	59.0 (32.7)
Screen (%)	67	21.3 [19.5, 23.1]	7.4	20.9 (7.8)
Screen among valid (%)	67	45.9 [41.9, 49.8]	16.2	48.6 (21.0)
Table (%)	67	28.8 [25.0, 32.7]	15.7	24.2 (27.5)
Table among valid (%)	67	54.1 [50.2, 58.1]	16.2	51.4 (21.0)

Table 1: Descriptive statistics (overall). Values are mean with 95% CI, SD, and median (IQR).

Position	Setup	Variable	n	Mean [95% CI]	SD	Median (IQR)
Sitting	Head-mounted	Unrelevant / Loss (%)	17	36.3 [31.0, 41.5]	10.2	33.1 (6.2)
Sitting	Head-mounted	Screen (%)	17	20.3 [15.0, 25.5]	10.2	23.1 (16.4)
Sitting	Head-mounted	Screen among valid (%)	17	30.6 [23.4, 37.7]	13.9	34.5 (19.3)
Sitting	Head-mounted	Table (%)	17	43.5 [39.4, 47.6]	8.0	42.2 (12.8)
Sitting	Head-mounted	Table among valid (%)	17	69.4 [62.3, 76.6]	13.9	65.5 (19.3)
Sitting	Remote	Unrelevant / Loss (%)	17	67.6 [64.7, 70.6]	5.7	66.6 (5.2)
Sitting	Remote	Screen (%)	17	19.7 [17.9, 21.5]	3.6	20.5 (4.5)
Sitting	Remote	Screen among valid (%)	17	61.6 [56.3, 66.9]	10.3	61.0 (15.1)
Sitting	Remote	Table (%)	17	12.7 [10.3, 15.1]	4.7	12.4 (6.3)
Sitting	Remote	Table among valid (%)	17	38.4 [33.1, 43.7]	10.3	39.0 (15.1)
Standing	Head-mounted	Unrelevant / Loss (%)	16	30.3 [25.8, 34.9]	8.6	29.0 (9.0)
Standing	Head-mounted	Screen (%)	16	26.2 [21.5, 30.8]	8.7	26.0 (7.7)
Standing	Head-mounted	Screen among valid (%)	16	36.9 [31.6, 42.2]	9.9	36.9 (10.8)
Standing	Head-mounted	Table (%)	16	43.5 [40.2, 46.8]	6.1	43.5 (8.0)
Standing	Head-mounted	Table among valid (%)	16	63.1 [57.8, 68.4]	9.9	63.1 (10.8)
Standing	Remote	Unrelevant / Loss (%)	17	64.3 [62.0, 66.5]	4.4	63.8 (4.0)
Standing	Remote	Screen (%)	17	19.2 [17.8, 20.6]	2.8	18.6 (3.5)
Standing	Remote	Screen among valid (%)	17	54.0 [51.0, 57.0]	5.8	53.9 (5.6)
Standing	Remote	Table (%)	17	16.5 [14.8, 18.3]	3.4	17.0 (4.2)
Standing	Remote	Table among valid (%)	17	46.0 [43.0, 49.0]	5.8	46.1 (5.6)

Table 2: Descriptive statistics by position and setup.

1 Tables

All tables below are auto-generated from the analysis CSV outputs.

Position	Setup	Calibration	Variable	n	Mean [95% CI]	SD	Median (IQR)
Sitting	Head-mounted	No issue	Unrelevant / Loss (%)	8	31.8 [26.8, 36.9]	6.0	32.5 (4.3)
Sitting	Head-mounted	Severe issues	Unrelevant / Loss (%)	4	48.9 [27.1, 70.7]	13.7	52.1 (8.3)
Sitting	Head-mounted	Slight issues	Unrelevant / Loss (%)	5	33.3 [30.5, 36.0]	2.2	33.3 (0.4)
Sitting	Remote	No issue	Unrelevant / Loss (%)	5	65.1 [56.1, 74.2]	7.3	63.0 (10.7)
Sitting	Remote	Severe issues	Unrelevant / Loss (%)	6	69.6 [65.1, 74.0]	4.3	68.4 (3.4)
Sitting	Remote	Slight issues	Unrelevant / Loss (%)	6	67.7 [61.6, 73.8]	5.8	66.0 (3.0)
Standing	Head-mounted	No issue	Unrelevant / Loss (%)	11	29.0 [24.7, 33.3]	6.3	29.3 (8.2)
Standing	Head-mounted	Severe issues	Unrelevant / Loss (%)	1	43.7	0.0	43.7 (0.0)
Standing	Head-mounted	Slight issues	Unrelevant / Loss (%)	4	30.6 [10.1, 51.2]	12.9	25.5 (11.7)
Standing	Remote	No issue	Unrelevant / Loss (%)	9	63.5 [59.7, 67.4]	5.0	63.1 (4.0)
Standing	Remote	Severe issues	Unrelevant / Loss (%)	3	63.7 [52.2, 75.3]	4.7	61.9 (4.4)
Standing	Remote	Slight issues	Unrelevant / Loss (%)	5	65.9 [61.9, 69.9]	3.2	64.5 (1.0)

Table 3: Descriptive statistics for data loss by position, setup, and calibration (loss label).

Position	Setup	Scene	Calibration	Variable	n	Mean [95% CI]	SD	Median (IQR)
Sitting	Head-mounted	Screen	No issue	Value	8	25.0 [19.4, 30.6]	6.7	26.1 (6.5)
Sitting	Head-mounted	Screen	Severe issues	Value	4	13.1 [−5.2, 31.5]	11.5	11.6 (11.5)
Sitting	Head-mounted	Screen	Slight issues	Value	5	18.5 [4.0, 32.9]	11.6	18.4 (14.5)
Sitting	Head-mounted	Table	No issue	Value	8	43.2 [38.7, 47.8]	5.4	43.3 (8.8)
Sitting	Head-mounted	Table	Severe issues	Value	4	38.0 [32.1, 43.9]	3.7	38.2 (3.7)
Sitting	Head-mounted	Table	Slight issues	Value	5	48.3 [34.0, 62.6]	11.5	51.4 (11.5)
Sitting	Remote	Screen	No issue	Value	12	19.8 [17.6, 22.0]	3.4	20.4 (3.2)
Sitting	Remote	Screen	Slight issues	Value	5	19.5 [14.1, 24.8]	4.3	22.2 (7.0)
Sitting	Remote	Table	No issue	Value	5	14.4 [7.1, 21.7]	5.8	14.5 (8.6)
Sitting	Remote	Table	Severe issues	Value	6	9.7 [5.7, 13.7]	3.8	9.8 (4.3)
Sitting	Remote	Table	Slight issues	Value	6	14.3 [10.5, 18.0]	3.5	14.7 (2.7)
Standing	Head-mounted	Screen	No issue	Value	11	27.0 [21.1, 33.0]	8.9	26.3 (7.1)
Standing	Head-mounted	Screen	Severe issues	Value	1	24.3	0.0	24.3 (0.0)
Standing	Head-mounted	Screen	Slight issues	Value	4	24.2 [7.4, 41.0]	10.6	26.9 (10.9)
Standing	Head-mounted	Table	No issue	Value	11	44.0 [39.8, 48.2]	6.3	43.4 (9.8)
Standing	Head-mounted	Table	Severe issues	Value	1	32.0	0.0	32.0 (0.0)
Standing	Head-mounted	Table	Slight issues	Value	4	45.1 [40.0, 50.3]	3.2	45.9 (4.2)
Standing	Remote	Screen	No issue	Value	6	19.3 [16.3, 22.3]	2.8	18.8 (3.9)
Standing	Remote	Screen	Severe issues	Value	1	20.9	0.0	20.9 (0.0)
Standing	Remote	Screen	Slight issues	Value	10	19.0 [16.8, 21.1]	3.0	18.2 (3.2)
Standing	Remote	Table	No issue	Value	9	17.8 [14.9, 20.8]	3.8	17.8 (2.8)
Standing	Remote	Table	Severe issues	Value	3	14.5 [8.9, 20.1]	2.3	13.7 (2.1)
Standing	Remote	Table	Slight issues	Value	5	15.4 [12.6, 18.3]	2.3	17.0 (4.2)

Table 4: Descriptive statistics by position, setup, scene, and scene-specific calibration.

DV	Subset	Factor	Test	Stat	p	p(BH)	Effect
Unrelevant / Loss (%)	Sitting	setup	mannwhitneyu	287.000	$< .001$	$< .001$	
Unrelevant / Loss (%)	Standing	setup	mannwhitneyu	272.000	$< .001$	$< .001$	
Table (%)	Sitting	setup	welch_t	-13.717	$< .001$	$< .001$	$g = -4.5$
Table (%)	Standing	setup	mannwhitneyu	0.000	$< .001$	$< .001$	
Screen (%)	Sitting	setup	mannwhitneyu	125.000	0.513	0.615	
Screen (%)	Standing	setup	mannwhitneyu	55.000	0.004	0.013	
Unrelevant / Loss (%)	Sitting / Remote	calibration_loss	anova_oneway	0.804	0.467	0.601	$\eta^2 = 0.1$
Unrelevant / Loss (%)	Sitting / Head-mounted	calibration_loss	anova_oneway	7.012	0.008	0.023	$\eta^2 = 0.5$
Unrelevant / Loss (%)	Standing / Remote	calibration_loss	kruskal	2.914	0.233	0.407	$\epsilon^2 = 0.0$
Unrelevant / Loss (%)	Standing / Head-mounted	calibration_loss	kruskal	2.149	0.342	0.473	$\epsilon^2 = 0.0$
Table (%)	Sitting / Remote	calibration_table	anova_oneway	2.124	0.156	0.313	$\eta^2 = 0.2$
Table (%)	Sitting / Head-mounted	calibration_table	anova_oneway	2.127	0.156	0.313	$\eta^2 = 0.2$
Table (%)	Standing / Remote	calibration_table	kruskal	2.559	0.278	0.417	$\epsilon^2 = 0.0$
Table (%)	Standing / Head-mounted	calibration_table	kruskal	2.782	0.249	0.407	$\epsilon^2 = 0.0$
Screen (%)	Sitting / Remote	calibration_screen	anova_oneway	0.023	0.880	0.880	$\eta^2 = 0.0$
Screen (%)	Sitting / Head-mounted	calibration_screen	anova_oneway	2.205	0.147	0.313	$\eta^2 = 0.2$
Screen (%)	Standing / Remote	calibration_screen	kruskal	0.422	0.810	0.880	$\epsilon^2 = -0.$
Screen (%)	Standing / Head-mounted	calibration_screen	kruskal	0.280	0.869	0.880	$\epsilon^2 = -0.$

Table 5: Omnibus inferential tests. BH indicates Benjamini–Hochberg correction across omnibus tests.

DV	Subset	Contrast	Test	Stat	p	p(BH)	E
Unrelevant / Loss (%)	Sitting / Remote	no_issue vs severe_issues	welch_t	-1.205	0.272	0.553	g
Unrelevant / Loss (%)	Sitting / Remote	no_issue vs slight_issues	welch_t	-0.655	0.532	0.553	g
Unrelevant / Loss (%)	Sitting / Remote	slight_issues vs severe_issues	welch_t	-0.615	0.553	0.553	g
Unrelevant / Loss (%)	Sitting / Head-mounted	no_issue vs severe_issues	welch_t	-2.376	0.084	0.158	g
Unrelevant / Loss (%)	Sitting / Head-mounted	slight_issues vs severe_issues	welch_t	-2.259	0.106	0.158	g
Unrelevant / Loss (%)	Sitting / Head-mounted	no_issue vs slight_issues	welch_t	-0.606	0.559	0.559	g
Unrelevant / Loss (%)	Standing / Remote	no_issue vs slight_issues	mannwhitneyu	10.000	0.112	0.336	
Unrelevant / Loss (%)	Standing / Remote	slight_issues vs severe_issues	mannwhitneyu	11.000	0.393	0.589	
Unrelevant / Loss (%)	Standing / Remote	no_issue vs severe_issues	mannwhitneyu	12.000	0.864	0.864	
Unrelevant / Loss (%)	Standing / Head-mounted	no_issue vs severe_issues	mannwhitneyu	0.000	0.167	0.500	
Unrelevant / Loss (%)	Standing / Head-mounted	no_issue vs slight_issues	mannwhitneyu	26.000	0.661	0.800	
Unrelevant / Loss (%)	Standing / Head-mounted	slight_issues vs severe_issues	mannwhitneyu	1.000	0.800	0.800	
Table (%)	Sitting / Remote	slight_issues vs severe_issues	welch_t	2.134	0.059	0.176	g
Table (%)	Sitting / Remote	no_issue vs severe_issues	welch_t	1.541	0.169	0.254	g
Table (%)	Sitting / Remote	no_issue vs slight_issues	welch_t	0.048	0.963	0.963	g
Table (%)	Sitting / Head-mounted	no_issue vs severe_issues	welch_t	1.946	0.085	0.178	g
Table (%)	Sitting / Head-mounted	slight_issues vs severe_issues	welch_t	1.881	0.119	0.178	g
Table (%)	Sitting / Head-mounted	no_issue vs slight_issues	welch_t	-0.924	0.397	0.397	g
Table (%)	Standing / Remote	no_issue vs severe_issues	mannwhitneyu	21.000	0.209	0.447	
Table (%)	Standing / Remote	no_issue vs slight_issues	mannwhitneyu	31.000	0.298	0.447	
Table (%)	Standing / Remote	slight_issues vs severe_issues	mannwhitneyu	9.000	0.786	0.786	
Table (%)	Standing / Head-mounted	no_issue vs severe_issues	mannwhitneyu	11.000	0.167	0.500	
Table (%)	Standing / Head-mounted	slight_issues vs severe_issues	mannwhitneyu	4.000	0.400	0.600	
Table (%)	Standing / Head-mounted	no_issue vs slight_issues	mannwhitneyu	19.000	0.753	0.753	
Screen (%)	Sitting / Head-mounted	no_issue vs severe_issues	welch_t	1.895	0.130	0.391	g
Screen (%)	Sitting / Head-mounted	no_issue vs slight_issues	welch_t	1.136	0.302	0.452	g
Screen (%)	Sitting / Head-mounted	slight_issues vs severe_issues	welch_t	0.685	0.517	0.517	g
Screen (%)	Standing / Remote	slight_issues vs severe_issues	mannwhitneyu	3.000	0.727	0.875	
Screen (%)	Standing / Remote	no_issue vs severe_issues	mannwhitneyu	2.000	0.857	0.875	
Screen (%)	Standing / Remote	no_issue vs slight_issues	mannwhitneyu	32.000	0.875	0.875	
Screen (%)	Standing / Head-mounted	no_issue vs slight_issues	mannwhitneyu	25.000	0.753	1.000	
Screen (%)	Standing / Head-mounted	no_issue vs severe_issues	mannwhitneyu	7.000	0.833	1.000	
Screen (%)	Standing / Head-mounted	slight_issues vs severe_issues	mannwhitneyu	2.000	1.000	1.000	

Table 6: Pairwise comparisons. p(BH) is corrected within each comparison family.

DV	Analysis	Term	F	p	p(BH)	η_p^2
Unrelevant / Loss (%)	anova	C(position)	6.232	0.015	0.031	0.090
Unrelevant / Loss (%)	anova	C(setup)	311.048	< .001	< .001	0.832
Unrelevant / Loss (%)	anova	C(position):C(setup)	0.487	0.488	0.542	0.008
Unrelevant / Loss (%)	rank_anova	C(position)	9.479	0.003	0.010	0.131
Unrelevant / Loss (%)	rank_anova	C(setup)	198.648	< .001	< .001	0.759
Unrelevant / Loss (%)	rank_anova	C(position):C(setup)	0.007	0.934	0.934	0.000
Table (%)	anova	C(position)	1.904	0.172	0.235	0.029
Table (%)	anova	C(setup)	416.007	< .001	< .001	0.868
Table (%)	anova	C(position):C(setup)	1.804	0.184	0.240	0.028
Table (%)	rank_anova	C(position)	4.217	0.044	0.078	0.063
Table (%)	rank_anova	C(setup)	210.775	< .001	< .001	0.770
Table (%)	rank_anova	C(position):C(setup)	2.757	0.102	0.161	0.042
Screen (%)	anova	C(position)	2.390	0.127	0.182	0.037
Screen (%)	anova	C(setup)	4.632	0.035	0.066	0.068
Screen (%)	anova	C(position):C(setup)	3.445	0.068	0.114	0.052
Screen (%)	rank_anova	C(position)	1.260	0.266	0.332	0.020
Screen (%)	rank_anova	C(setup)	8.008	0.006	0.019	0.113
Screen (%)	rank_anova	C(position):C(setup)	2.413	0.125	0.182	0.037
Table among valid (%)	anova	C(position)	0.086	0.771	0.798	0.001
Table among valid (%)	anova	C(setup)	90.519	< .001	< .001	0.590
Table among valid (%)	anova	C(position):C(setup)	7.591	0.008	0.019	0.108
Table among valid (%)	rank_anova	C(position)	0.554	0.459	0.530	0.009
Table among valid (%)	rank_anova	C(setup)	132.274	< .001	< .001	0.677
Table among valid (%)	rank_anova	C(position):C(setup)	6.202	0.015	0.031	0.090
Screen among valid (%)	anova	C(position)	0.086	0.771	0.798	0.001
Screen among valid (%)	anova	C(setup)	90.519	< .001	< .001	0.590
Screen among valid (%)	anova	C(position):C(setup)	7.591	0.008	0.019	0.108
Screen among valid (%)	rank_anova	C(position)	0.554	0.459	0.530	0.009
Screen among valid (%)	rank_anova	C(setup)	132.274	< .001	< .001	0.677
Screen among valid (%)	rank_anova	C(position):C(setup)	6.202	0.015	0.031	0.090

Table 7: Factorial analysis (setup × position). Partial η_p^2 reported.

DV	Contrast	Test	Stat	p	p(BH)	Effe
Unrelevant / Loss (%)	sitting__head_mounted vs sitting__remote	mannwhitneyu	2.000	< .001	< .001	
Unrelevant / Loss (%)	sitting__remote vs standing__head_mounted	mannwhitneyu	272.000	< .001	< .001	
Unrelevant / Loss (%)	standing__head_mounted vs standing__remote	mannwhitneyu	0.000	< .001	< .001	
Unrelevant / Loss (%)	sitting__head_mounted vs standing__remote	mannwhitneyu	5.000	< .001	< .001	
Unrelevant / Loss (%)	sitting__head_mounted vs standing__head_mounted	mannwhitneyu	194.000	0.038	0.046	
Unrelevant / Loss (%)	sitting__remote vs standing__remote	mannwhitneyu	203.000	0.046	0.046	
Table (%)	sitting__head_mounted vs sitting__remote	mannwhitneyu	289.000	< .001	< .001	
Table (%)	sitting__head_mounted vs standing__remote	mannwhitneyu	289.000	< .001	< .001	
Table (%)	sitting__remote vs standing__head_mounted	mannwhitneyu	0.000	< .001	< .001	
Table (%)	standing__head_mounted vs standing__remote	mannwhitneyu	272.000	< .001	< .001	
Table (%)	sitting__remote vs standing__remote	mannwhitneyu	72.000	0.013	0.016	
Table (%)	sitting__head_mounted vs standing__head_mounted	mannwhitneyu	129.000	0.815	0.815	
Screen (%)	sitting__remote vs standing__head_mounted	mannwhitneyu	53.000	0.003	0.011	
Screen (%)	standing__head_mounted vs standing__remote	mannwhitneyu	217.000	0.004	0.011	
Screen (%)	sitting__head_mounted vs standing__head_mounted	mannwhitneyu	98.000	0.177	0.354	
Screen (%)	sitting__head_mounted vs standing__remote	mannwhitneyu	172.000	0.352	0.529	
Screen (%)	sitting__head_mounted vs sitting__remote	mannwhitneyu	164.000	0.513	0.615	
Screen (%)	sitting__remote vs standing__remote	mannwhitneyu	157.000	0.679	0.679	
Table among valid (%)	sitting__head_mounted vs sitting__remote	mannwhitneyu	285.000	< .001	< .001	
Table among valid (%)	sitting__head_mounted vs standing__remote	mannwhitneyu	282.000	< .001	< .001	
Table among valid (%)	sitting__remote vs standing__head_mounted	mannwhitneyu	9.000	< .001	< .001	
Table among valid (%)	standing__head_mounted vs standing__remote	mannwhitneyu	255.000	< .001	< .001	
Table among valid (%)	sitting__remote vs standing__remote	mannwhitneyu	76.000	0.019	0.023	
Table among valid (%)	sitting__head_mounted vs standing__head_mounted	mannwhitneyu	168.000	0.257	0.257	
Screen among valid (%)	sitting__head_mounted vs sitting__remote	mannwhitneyu	4.000	< .001	< .001	
Screen among valid (%)	sitting__head_mounted vs standing__remote	mannwhitneyu	7.000	< .001	< .001	
Screen among valid (%)	sitting__remote vs standing__head_mounted	mannwhitneyu	263.000	< .001	< .001	
Screen among valid (%)	standing__head_mounted vs standing__remote	mannwhitneyu	17.000	< .001	< .001	
Screen among valid (%)	sitting__remote vs standing__remote	mannwhitneyu	213.000	0.019	0.023	
Screen among valid (%)	sitting__head_mounted vs standing__head_mounted	mannwhitneyu	104.000	0.257	0.257	

Table 8: Post-hoc comparisons between position × setup cells (BH corrected).

2 Figures

Figures below are generated by the plotting scripts in `analysis/plots/`.

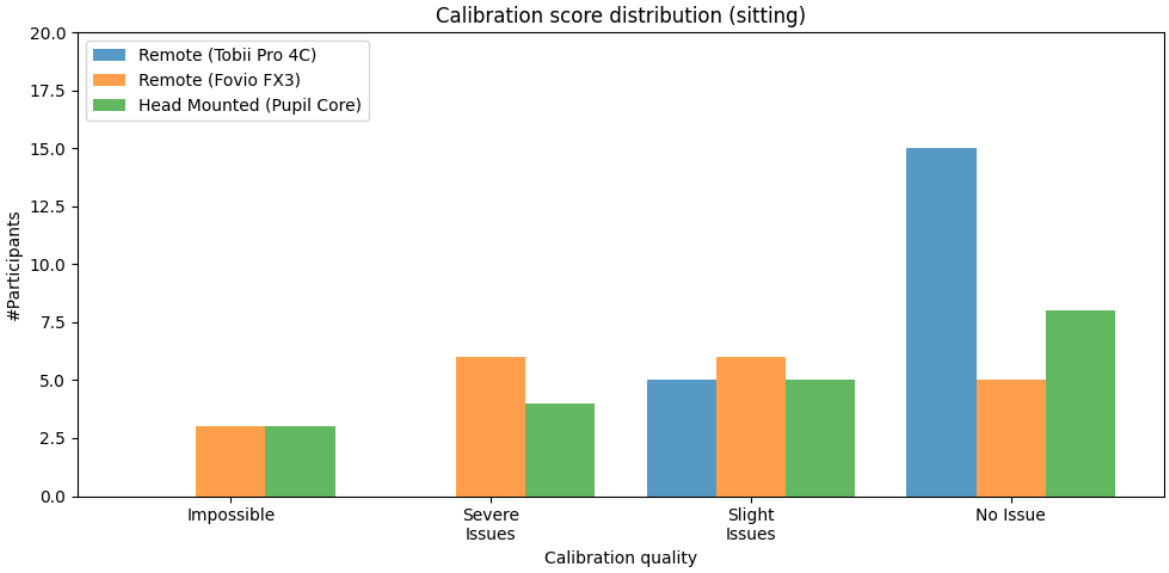

Figure 1: calibration_score_sitting.png

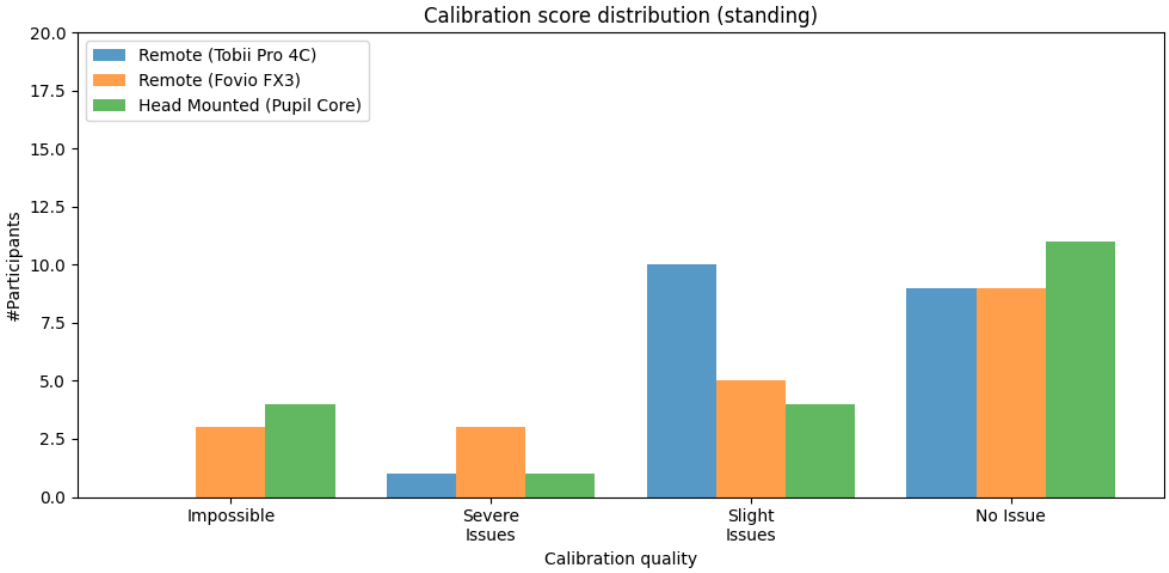

Figure 2: calibration_score_standing.png

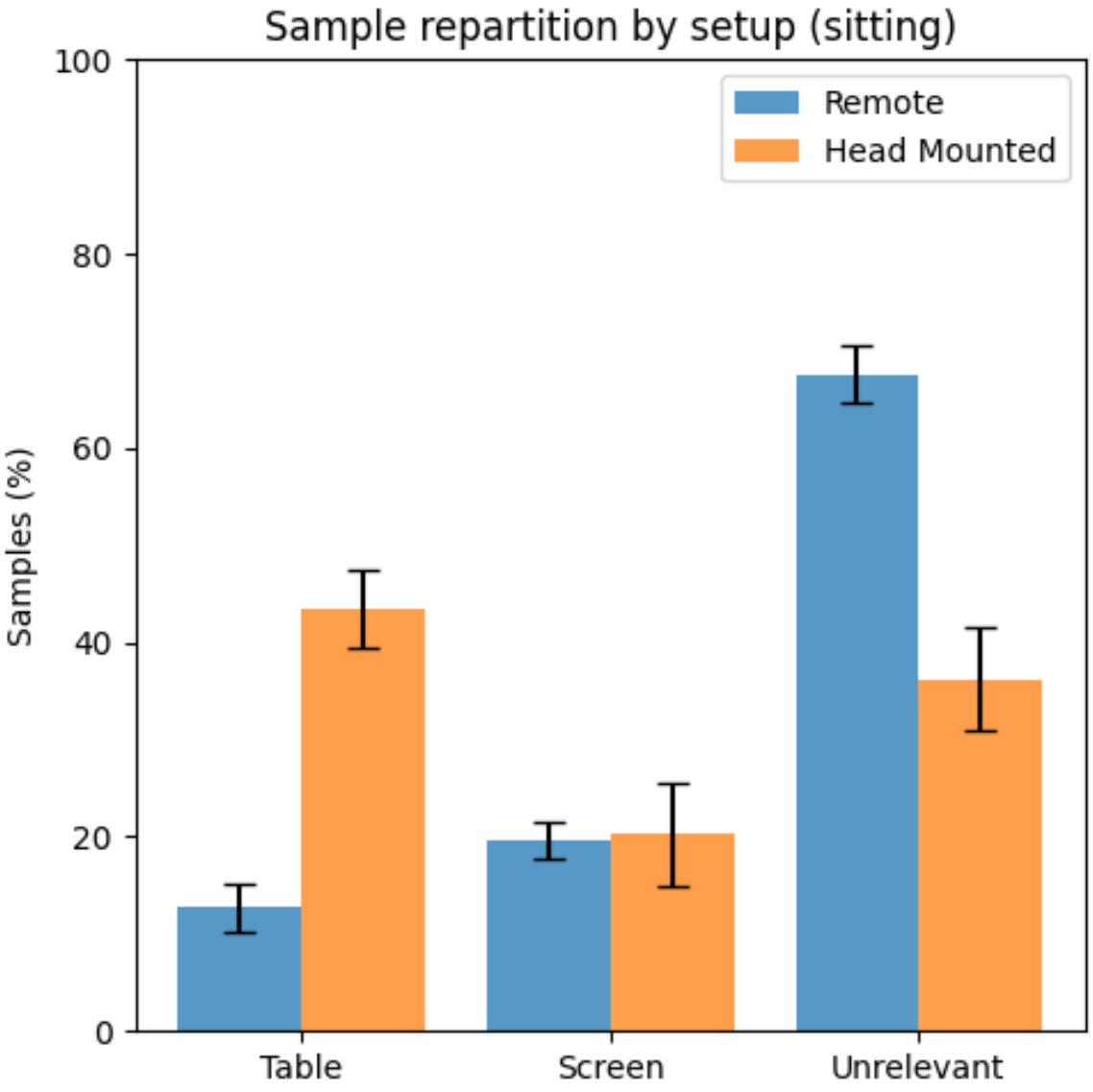

Figure 3: data_loss_sitting.png

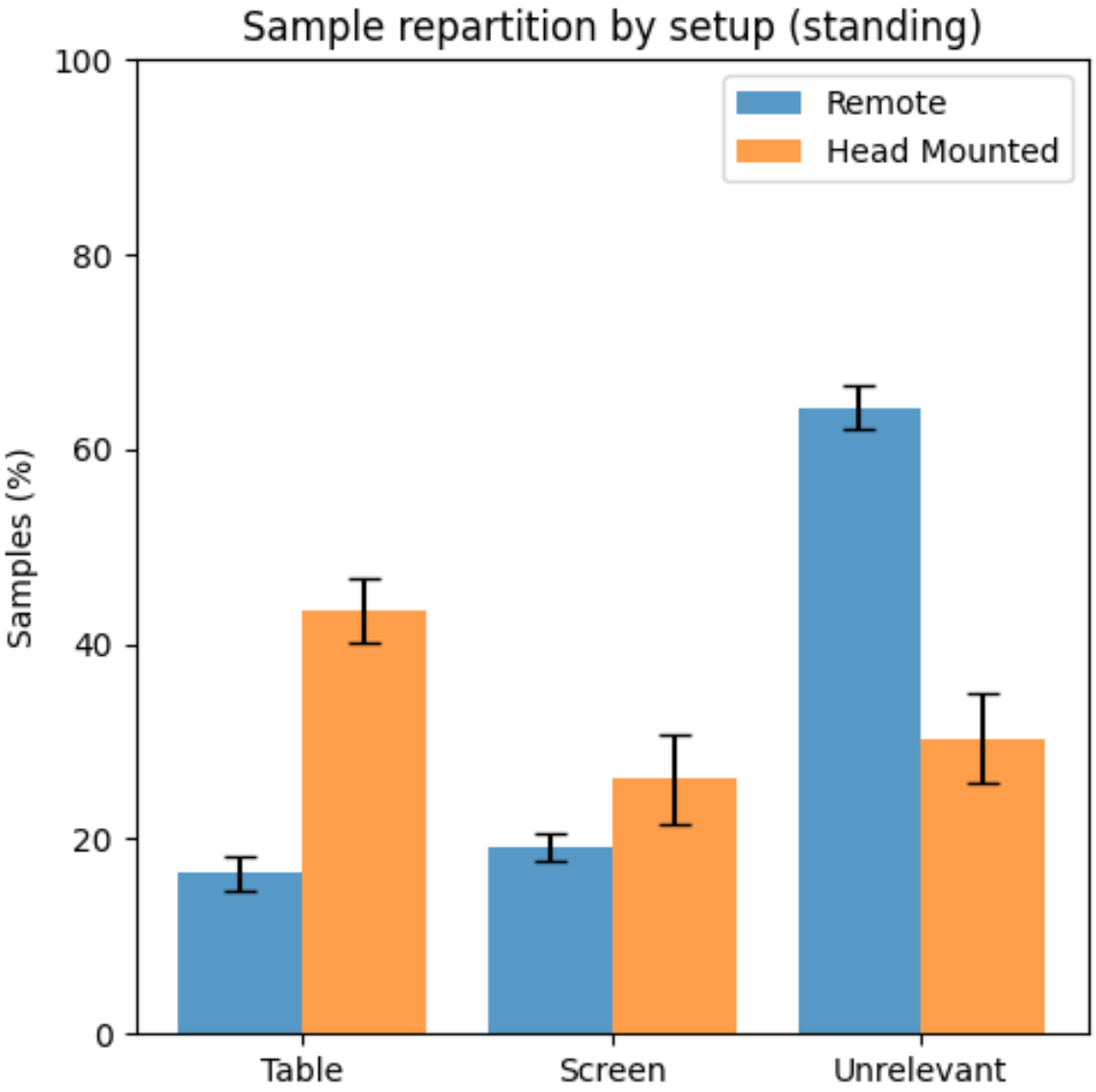

Figure 4: data_loss_standing.png

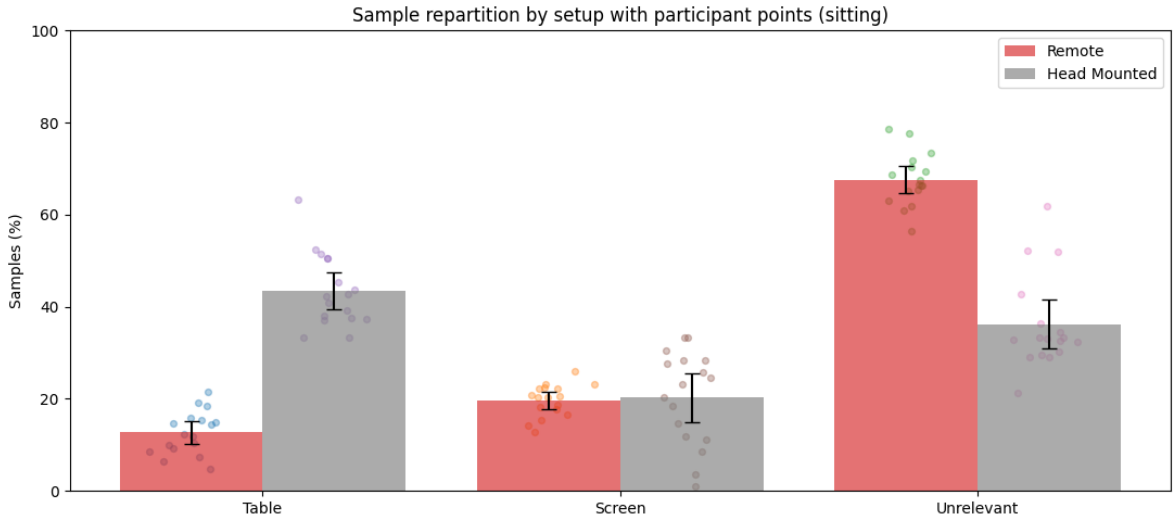

Figure 5: data_loss_points_sitting.png

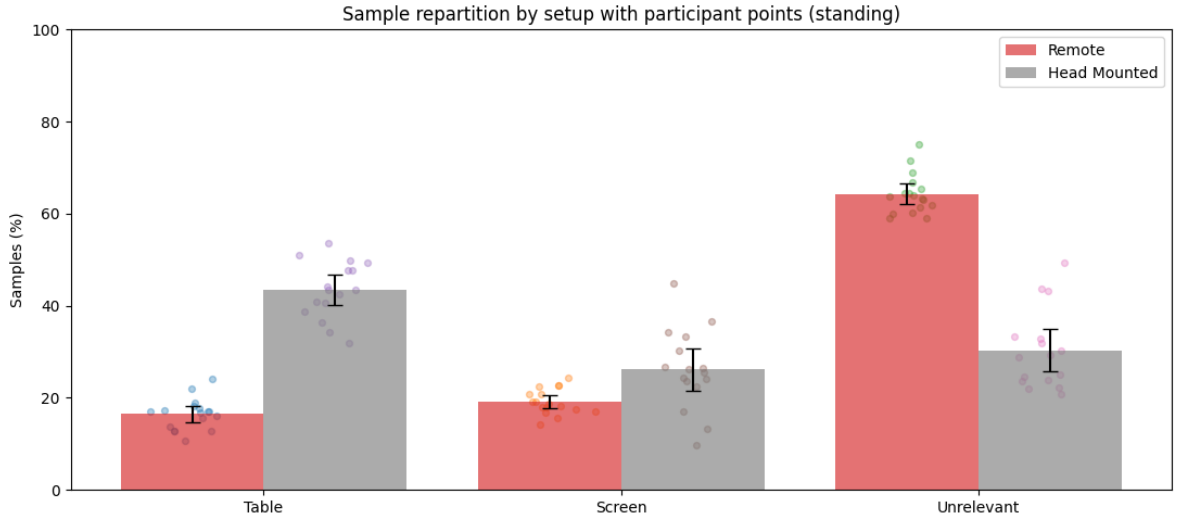

Figure 6: data_loss_points_standing.png

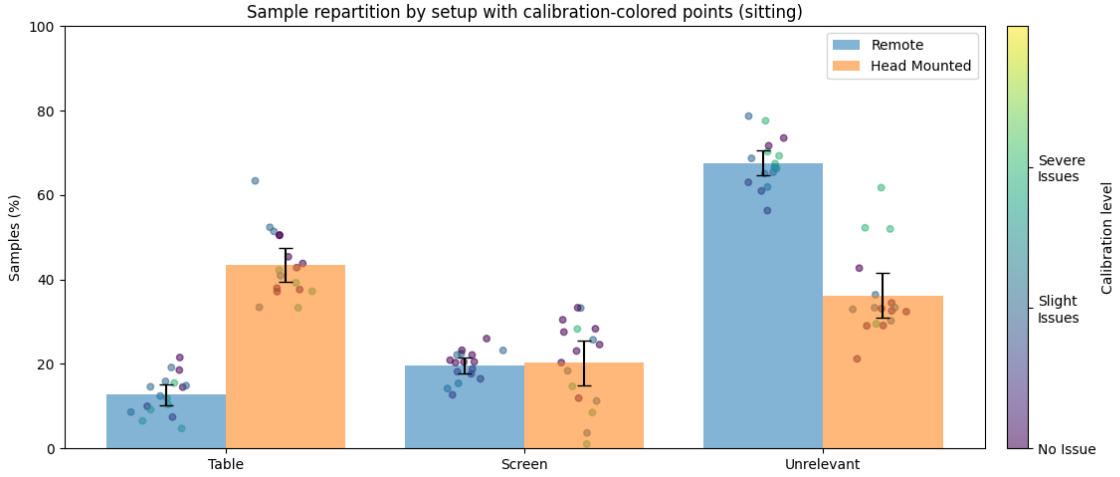

Figure 7: data_loss_points_by_calib_sitting.png

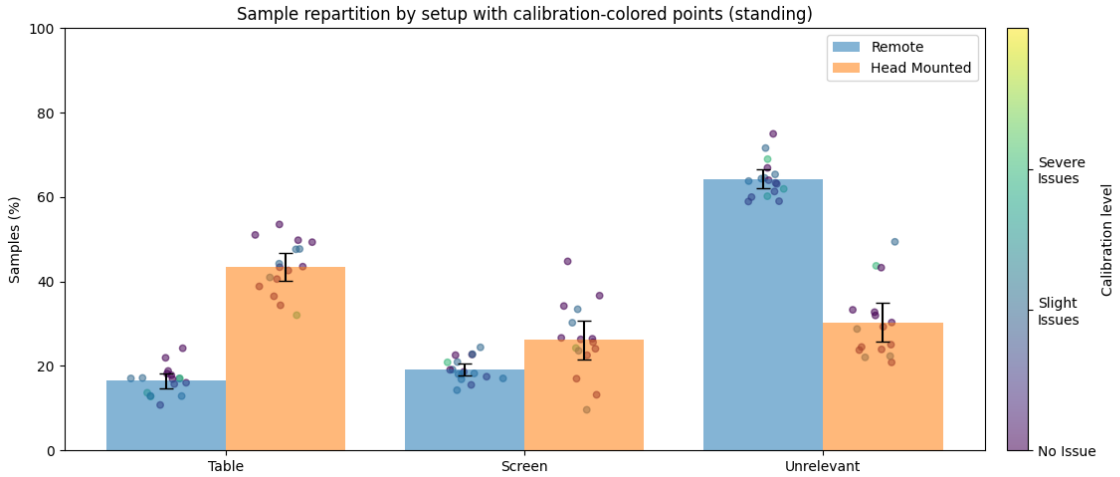

Figure 8: data_loss_points_by_calib_standing.png

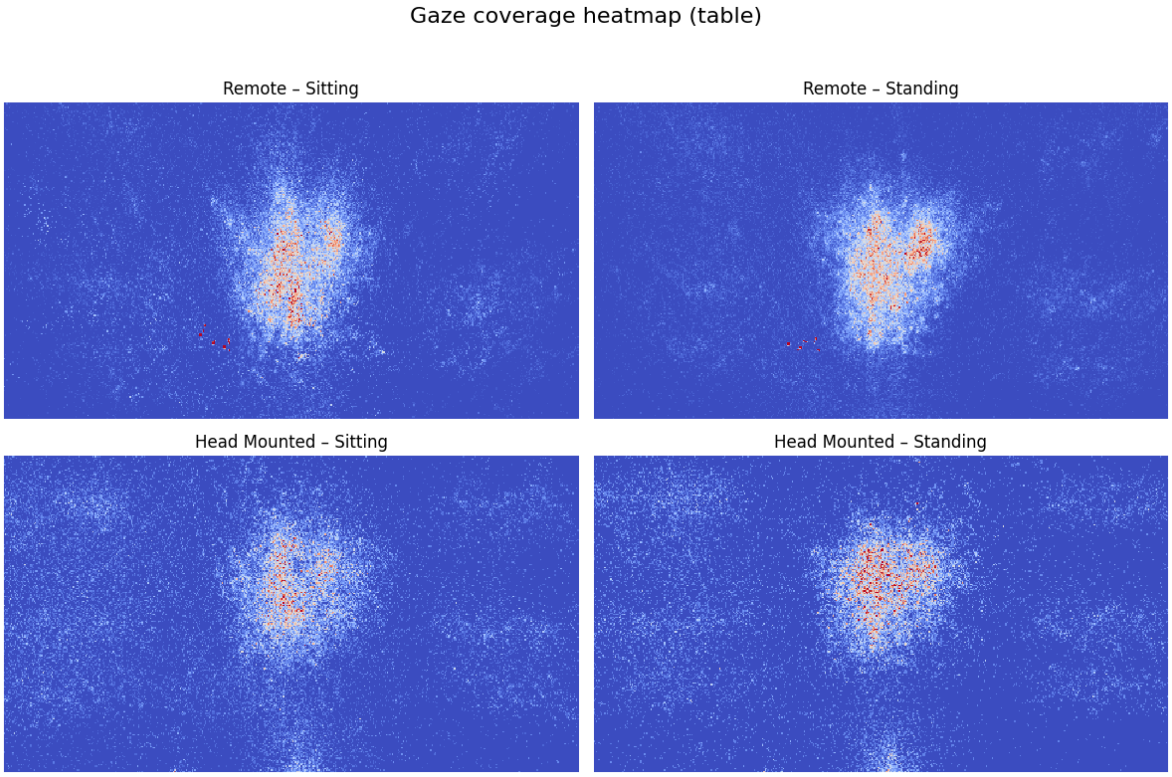

Figure 9: heatmap_table.png

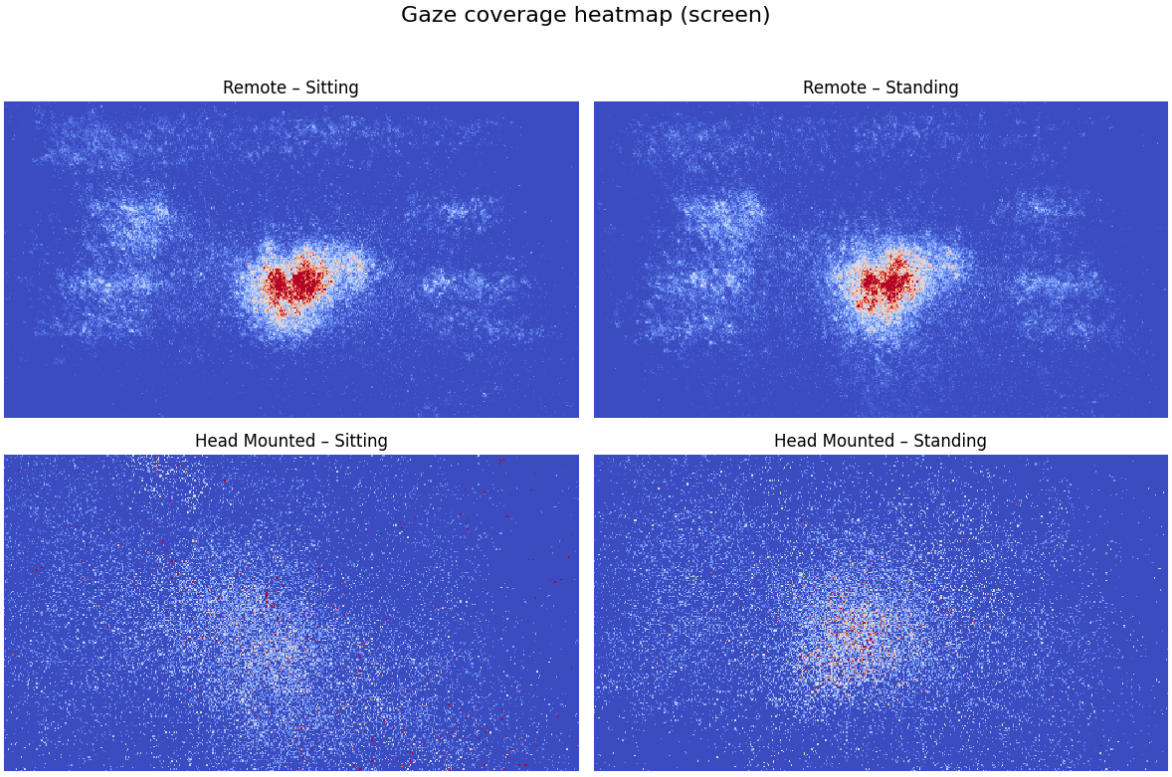

Figure 10: heatmap_screen.png

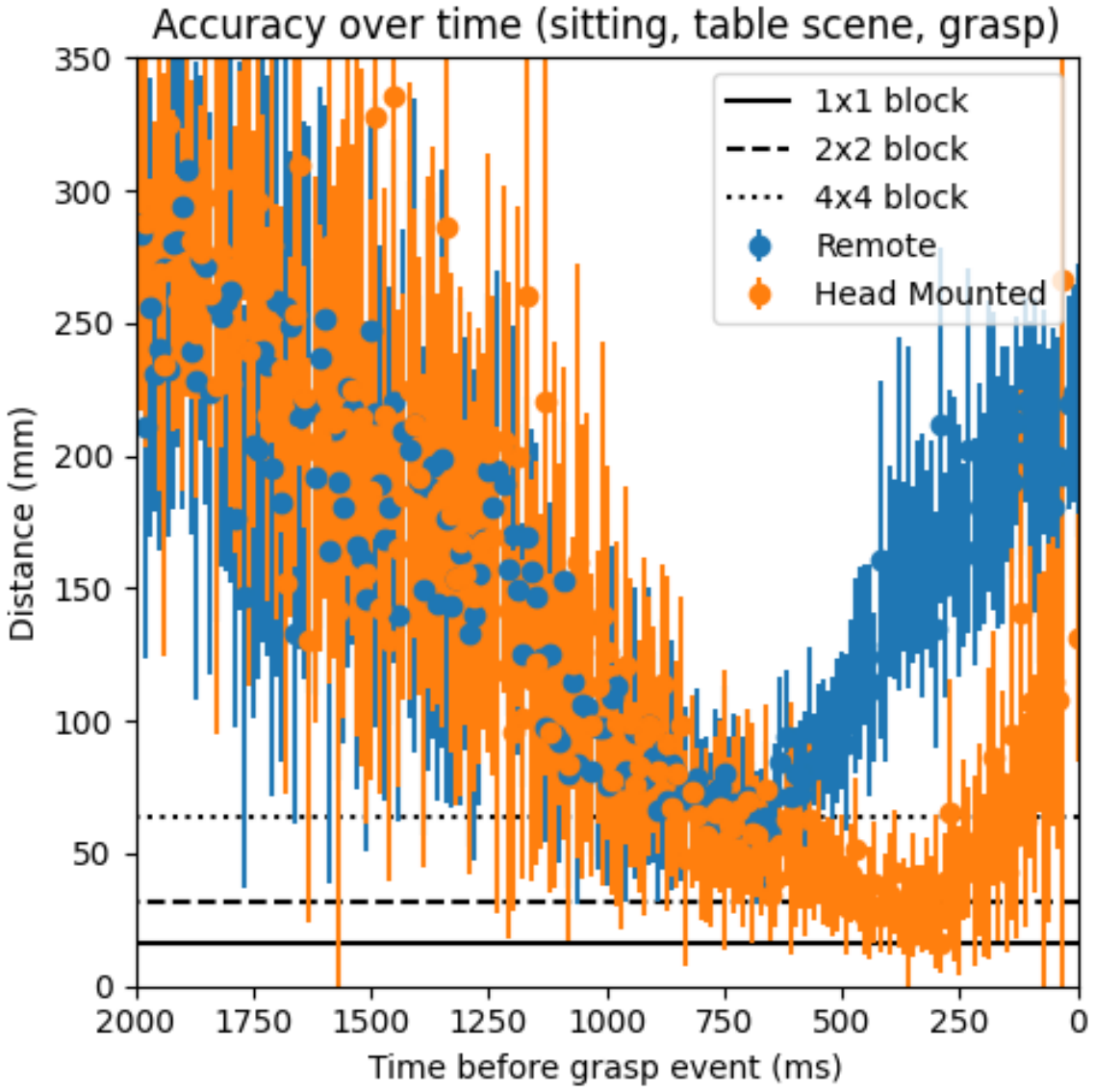

Figure 11: accuracy_time_sitting_table_grasp.png

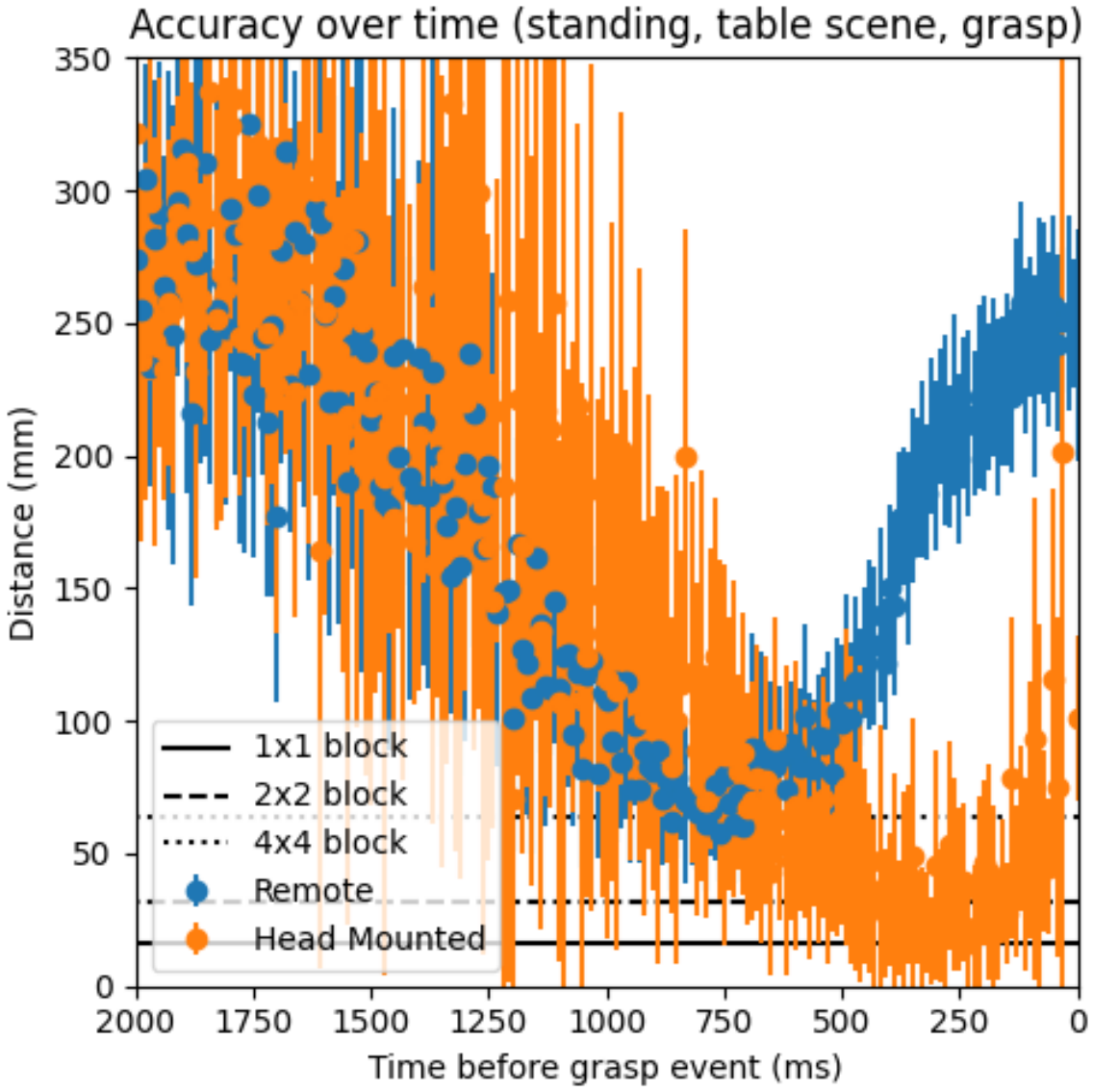

Figure 12: accuracy_time_standing_table_grasp.png

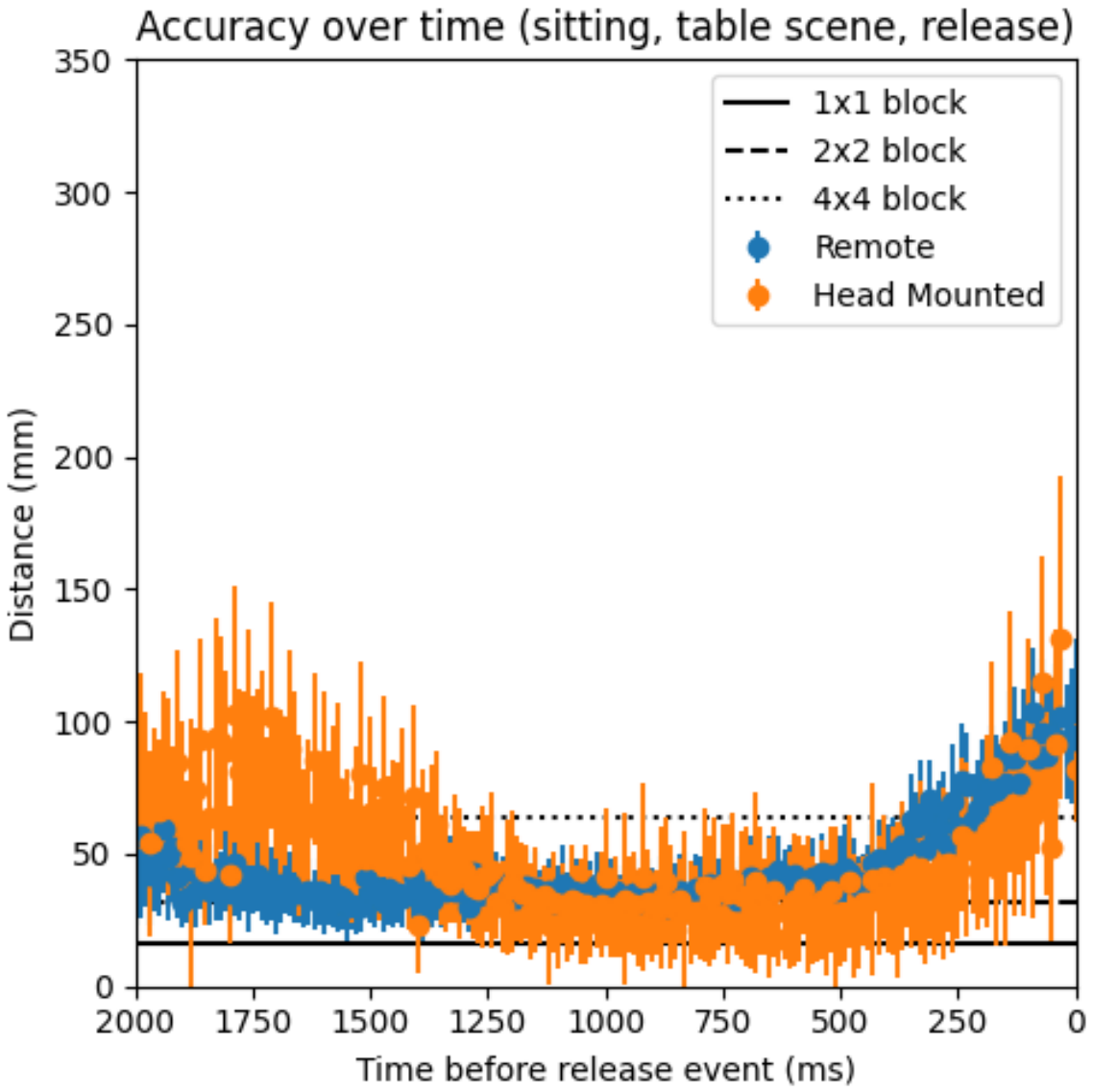

Figure 13: accuracy_time_sitting_table_release.png

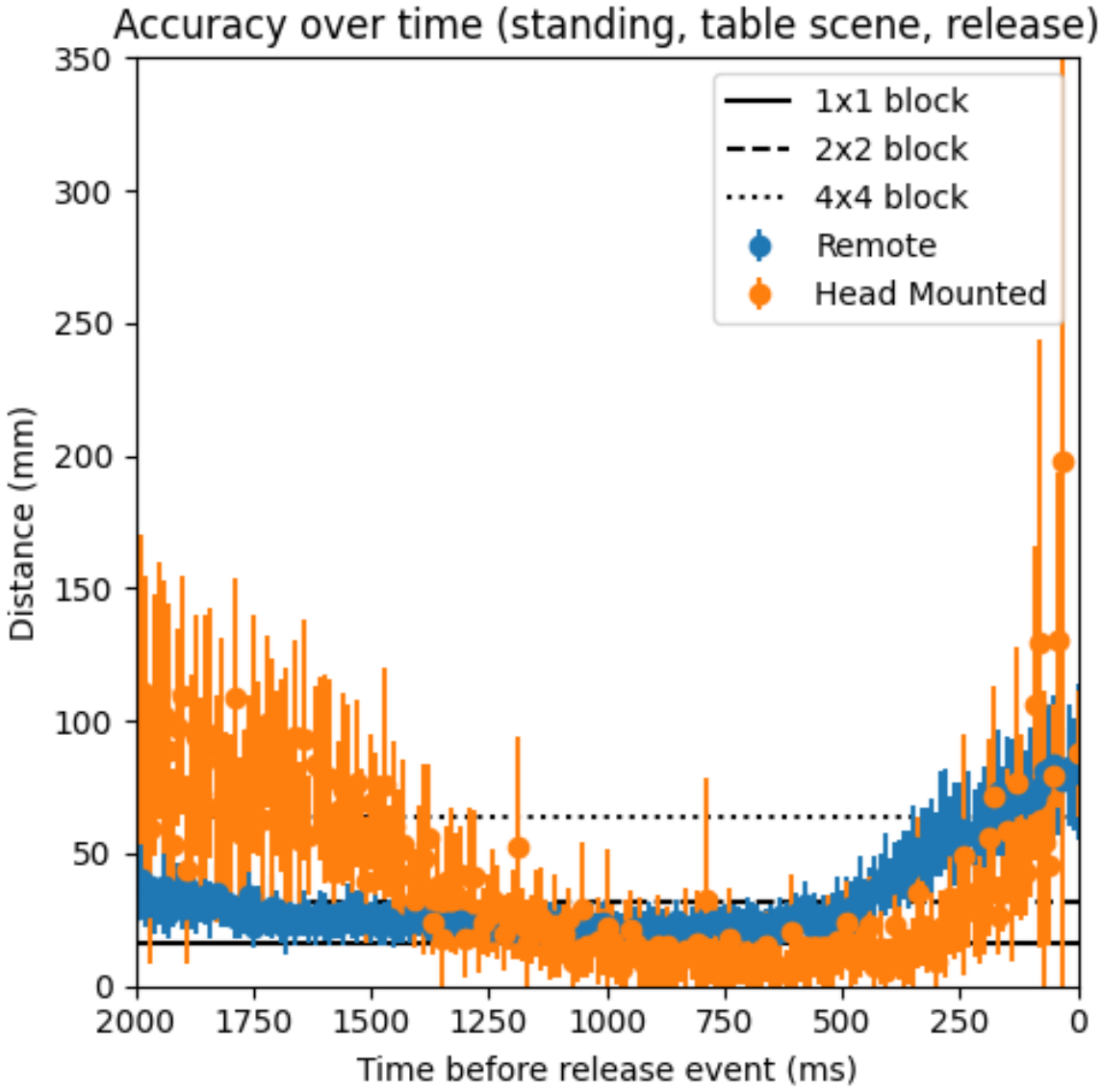

Figure 14: accuracy_time_standing_table_release.png