

Inferential Eye Movement Control while Following Dynamic Gaze

Nicole X. Han

Department of Psychological and Brain Sciences

University of California, Santa Barbara

Institute for Collaborative Biotechnologies

University of California, Santa Barbara

xhan01@ucsb.edu

Miguel P. Eckstein

Department of Psychological and Brain Sciences

Department of Electrical and Computer Engineering

Department of Computer Science

Institute for Collaborative Biotechnologies

University of California, Santa Barbara

migueleckstein@ucsb.edu

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71

ABSTRACT

Attending to other people’s gaze is evolutionarily important to make inferences about intentions and actions. Gaze influences covert attention and triggers eye movements. However, we know little about how the brain controls the fine-grain dynamics of eye movements during gaze following. Observers followed people’s gaze shifts in videos during search and we related the observer eye movement dynamics to the timecourse of gazer head movements extracted by a deep neural network. We show that the observers’ brains use information in the visual periphery to execute predictive saccades that anticipate the information in the gazer’s head direction by 190-350 ms. The brain simultaneously monitors moment-to-moment changes in the gazer’s head velocity to dynamically alter eye movements and re-fixate the gazer (reverse saccades) when the head accelerates before the initiation of the first forward gaze-following saccade. Using saccade-contingent manipulations of the videos, we experimentally show that the reverse saccades are planned concurrently with the first forward gaze-following saccade and have a functional role in reducing subsequent errors fixating on the gaze goal. Together, our findings characterize the inferential and functional nature of the fine-grain eye movement dynamics of social attention.

72 Introduction

73 Eye movements are involved in almost every daily human activity, from searching for your
74 apartment key, identifying a friend, reading, and preparing a sandwich. People make about
75 three ballistic eye movements (saccades) per second orienting the central part of the vision (the
76 foveola) to regions of interest in the world and acquiring high acuity visual information (Bahill
77 et al., 1975). A foveated visual system allocates more retinal cells and cortical neurons to the
78 central part of the visual field and allows for computational and metabolic savings. But it
79 requires that eye movements are programmed intelligently to overcome the deficits of
80 peripheral processing. People execute eye movements effortlessly, rapidly, and automatically
81 giving the impression that these might seem fairly random. However, there are sophisticated
82 computations in the brain that control eye movements involved in fine spatial judgments (Intoy
83 & Rucci, 2020; Rucci et al., 2007), search (Araujo et al., 2001; Eckstein et al., 2015; Hoppe &
84 Rothkopf, 2019; Najemnik & Geisler, 2005), object identification (Renninger et al., 2007), face
85 recognition (Or et al., 2015; Peterson & Eckstein, 2012), reading (Legge et al., 1997, 2002) and
86 motor actions (Ballard et al., 1995; M. Hayhoe & Ballard, 2005).

87
88 Following the gaze of others (gaze-following) with eye movements is critical to infer others'
89 intentions, current interests, and future actions (Emery, 2000; Kleinke, 1986). Gaze-following
90 behavior can be observed as early as 8-10 months in infants and is widely found in nonhumans
91 such as macaques, dogs, and goats (Brooks & Meltzoff, 2005; Kaminski et al., 2005; Senju &
92 Csibra, 2008; Shepherd, 2010; Wallis et al., 2015). The ability to perceive others' gaze direction
93 accurately and plan eye movements is essential for infants to engage in social interactions to
94 learn objects and languages (Carpenter et al., 1998; Morales et al., 1998, 2000; Woodward,
95 2003). People are extremely sensitive to others' direction of gaze (Kleinke, 1986; Langton &
96 Bruce, 1999). When people observe someone's gaze, they orient covert attention and eye
97 movements toward the gazed location, which improves the detection of a target appearing in
98 the gaze direction (Driver et al., 1999; Egeth & Yantis, 1997; Friesen et al., 2004; Han & Eckstein,
99 2022; Jonides & Jonides, 1981; Kingstone et al., 2003; Mulckhuyse & Theeuwes, 2010). Gaze
100 cueing has also been proposed as an important correlate of autism spectrum disorder (Baron-
101 Cohen, 2001; Leekam et al., 1998; Nation & Penny, 2008) and important in child development
102 (Brooks & Meltzoff, 2005).

103
104 The majority of studies investigating gaze-following (but see Gregory, 2021; Han & Eckstein,
105 2022; Wang et al., 2014) use static images of the eyes or the face in isolation, which are far
106 from the more ecological real-world behaviors of individuals moving their heads and eyes when
107 orienting attention. That gaze cueing triggers eye movements is well known, but the dynamics
108 of eye movements when observing gaze behaviors with naturalistic dynamic stimuli are not
109 known. Studies have investigated how the brain integrates temporal information to program

110 saccades and how it integrates foveal and peripheral information (Stewart et al., 2020; Wolf et
111 al., 2022; Wolf & Schütz, 2015) but have relied on artificial or simplified stimuli.

112
113 Little is known about what features across the visual field influence eye movements during
114 gaze-following, their temporal dynamics, and their functionality. How does the brain rely on the
115 features of the gazer's head and peripheral visual information about likely gaze goals to
116 program eye movements? Do observers wait for the gazer's head movement to end before
117 initiating the first gaze-following saccades? Do visual properties of the gazer's head influence
118 the programming of eye movements? Answering these questions has been difficult because
119 they require a well-controlled real-world data set, moment-to-moment characterization of the
120 gazer's features, and experimental manipulations that alter peripheral information while
121 maintaining the gazer's information unaltered.

122
123 Here, we created a collection of in-house videos of dynamic gaze behaviors in real-world
124 settings by instructing actors to direct their gaze to specific people on the filming set (Figure
125 1a). We used digital editing tools to erase potential gaze goals while maintaining the gazer's
126 movements unaltered and preserving the video's background (Figure 1b-e). We then asked
127 experiment participants that watched the videos to follow the gaze shifts in the videos and
128 decide whether a specific target person was present or absent (Figure 1a).

129
130 Our first goal is to assess the impact of peripheral gaze-goal information on the saccade error
131 and timing. Second, we aimed to elucidate how the brain temporally processes visual
132 information to influence saccade programming during gaze-following. To extract features of the
133 videos that we hypothesized would influence saccade planning we used a state-of-the-art
134 artificial intelligence (AI) model (Chong et al., 2020) to make moment-to-moment estimates of
135 the gazer's head direction in the videos. We assessed how observers' saccade direction, timing,
136 and errors related to the extracted features to gain insight into the brain computations during
137 saccade planning.

138

139 **Results**

140 ***Integration of peripheral information to guide gaze-following saccades***

141 Twenty-five observers viewed 80 in-house videos (1.2s long, different settings) of an actor (gazer)
142 actively shifting his/her head and gaze to look at another person (gaze-goal) in the video.
143 Participants' initial fixation was on the gazer's head. They were instructed to look where the gazer
144 looks and report whether a specific target person was the gaze goal (Figure 1a). In 25 % of the
145 videos the target person was present and always the gaze goal (Figure 1b). In another 25 %, a
146 distractor person (Figure 1c) was the gaze goal and the target person was absent. In the

147 remaining 50%, no person was at the gazed location (d-e). The target/distractor absent videos
 148 were created by digitally removing the person at the gaze-goal location. The gazers' visual
 149 information in the videos was identical in the target/distractor present vs. absent videos (Figure
 150 1b vs. 1d and 1c vs. 1e). Throughout the trial, we measured eye position and detected the onset
 151 of saccades registered to the video timing. Observers typically executed 3-5 saccades. Figure 1f
 152 shows a histogram of the number of executed saccades per trial.

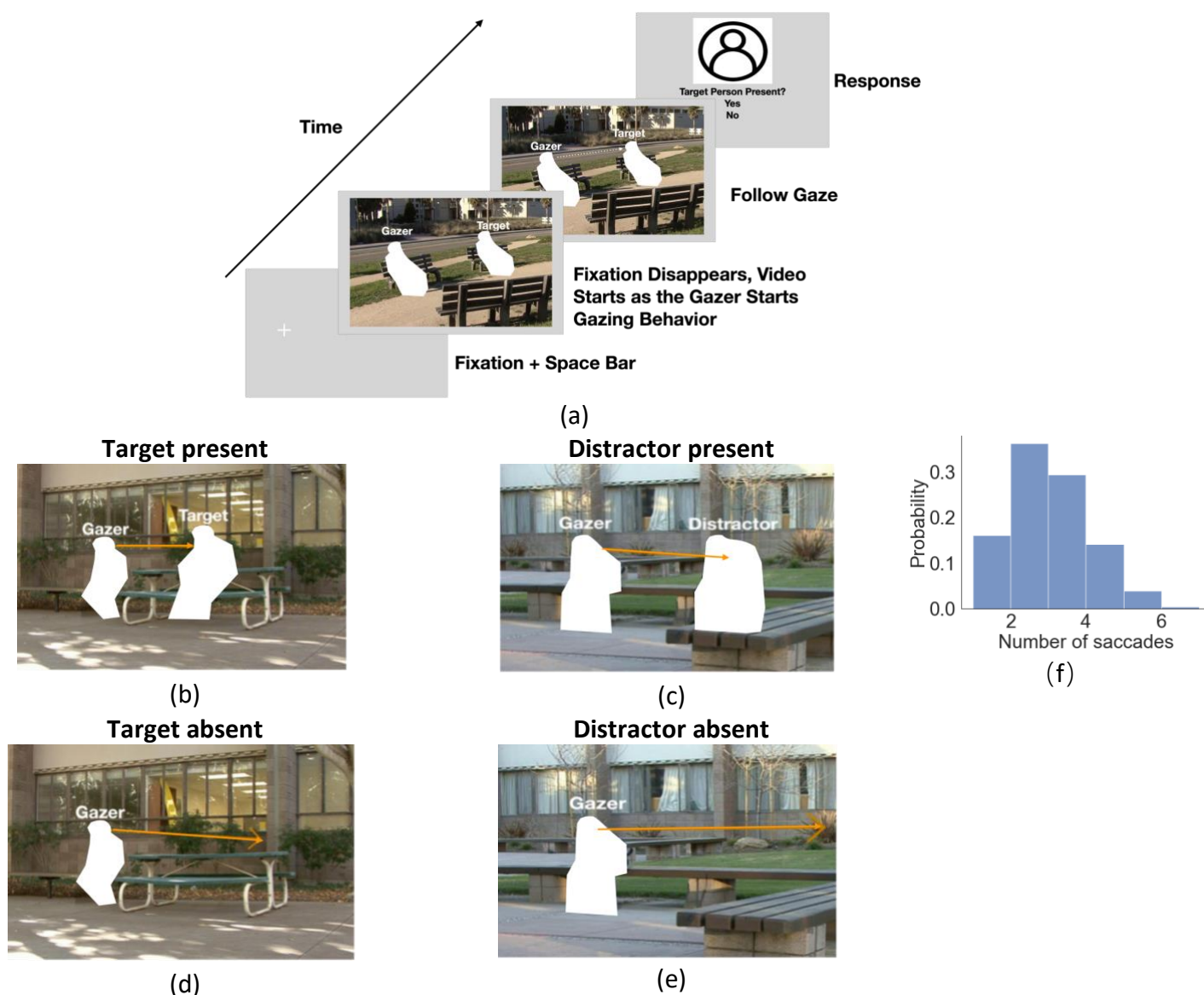


Figure 1. (a) Timeline for each trial. The participants fixated on the fixation cross and pressed the space bar to initialize the trial. The cross was located at the gazer's head, and the trial would not start if the eye fixation moved away from the cross by 1.5°. The cross disappeared as the video started with the gazer starting looking at the designated gazed person (50% present: 25% target and 25% distractor, 50% absent). Participants were instructed to follow the gaze direction and clicked to respond whether the target person was present or not. (b)-(e) Example video frames of the gazer looking at the gaze goal (distractor or target) either with the person present (b,c) or absent

(digitally deleted, d,e). The orange arrow vector is the gaze estimation from a deep neural network model, with details presented in the following section. Note that all the text annotations and arrows are just for illustration purposes and were not presented during experiments. (f) Histogram of the number of saccades participants executed per trial.

153

154 To investigate the effect of peripheral information on eye movement planning, we tested the
155 influence of the presence of a person at the gaze goal on the first saccade error and timing. Figure
156 2a-d show heat maps of first saccade endpoint distributions across all observers (for one
157 particular video) and illustrate how the peripheral presence of a person at the gaze goal reduces
158 the error of the first fixation. To quantify the error we calculated the mean Euclidean distance in
159 degrees of visual angle ($^{\circ}$) between the saccade endpoint and the center of the gazed person's
160 head. We found that the presence of a person at the gaze goal in the periphery reduced the
161 saccade error (Figure 2e, 2 (present or absent) \times 2 (target or distractor) ANOVA, $F(1,24)=259$,
162 $p=2.3e-14$). Saccade error was higher when a person was absent vs. present for both target (5.08°
163 vs. 2.50° , $p=1.4e-91$, post-hoc paired-test with False Discovery Rate, FDR) and distractor person
164 (5.25° vs. 2.60° , $p=6.8e-106$, FDR). The presence of a person at the gaze goal also impacts the
165 first saccade latency, ($F(1,24)=50.5$, $p=2.4e-07$). The saccade latency was significantly higher
166 (Figure 2f) when a person was absent vs. present at the gaze goal for both target ($0.37s$ vs. $0.31s$,
167 $p=1.4e-16$) and distractor ($0.38s$ vs. $0.31s$, $p=2.0e-19$) trials. There was no difference when the
168 target or distractor person was at the gaze-goal locations neither for first saccade error
169 ($F(1,24)=1.94$, $p=0.18$) nor first saccade latency ($F(1,24)=2.15$, $p=0.15$, Figure 2e-f).

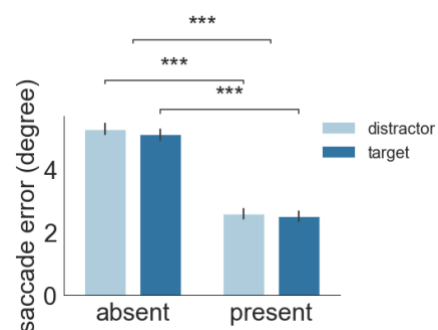
170



(a) Target-present



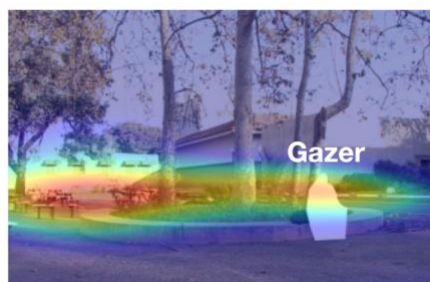
(b) Target-absent



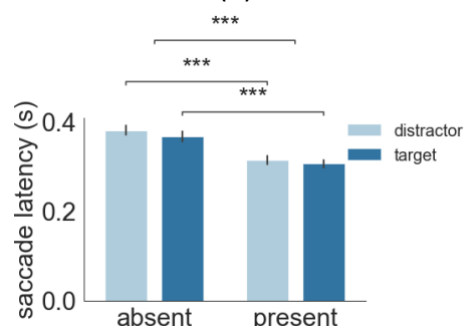
(e)



(c) Distractor-present



(d) Distractor-absent



(f)

Figure 2 (a)-(d) Examples of first gaze-following saccade endpoint density maps for target-present. (e) first gaze-following saccade endpoint error. (f) first gaze-following saccade latency.

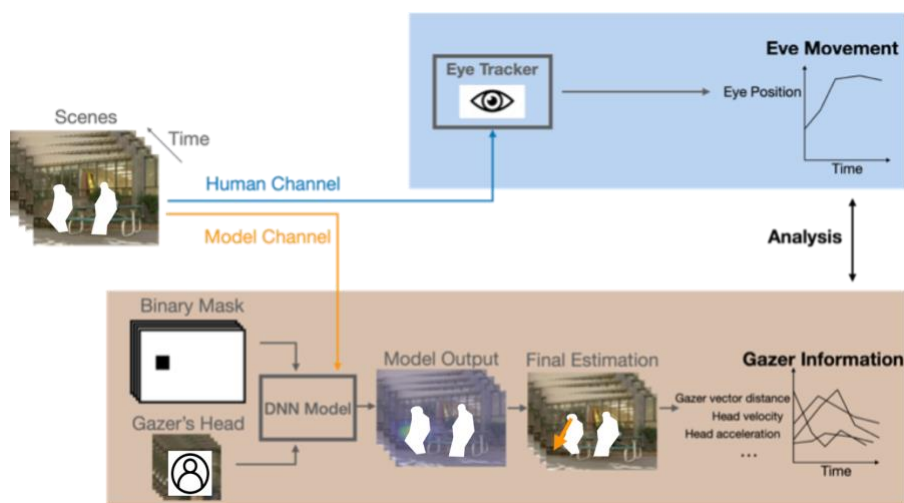
171

172

173 **Relating eye movement dynamics to gazer information**

174 To relate the dynamics of eye movements (Figure 3a) to the gazer's head information throughout
175 the video, we estimated gaze direction using a state-of-the-art deep neural network (DNN) model
176 (Chong et al., 2020, Figure 3a, see methods for details). The accuracy of the DNN model in
177 estimating the gaze goal location for these images is comparable to that of humans for
178 target/distractor present and superior to humans for absent trials (Han et al., 2021). For each
179 video frame, the model generated a gazer vector in which the start point was the gazer's eye
180 position, and the endpoint was the model estimated gaze-goal location. From the frame-to-frame
181 gazer vector, we calculated gazer vector distance in degrees of visual angle ($^{\circ}$), angular
182 displacement (deg), head velocity (deg/s), and head acceleration (deg/s²) at a sampling rate of
183 30 frames/sec (see Methods for detailed calculation, see Figure 3b gaze information definitions).
184 We could then relate the observers' saccade execution times to the moment-to-moment changes
185 in the gazer vector's measures. We also quantified, from the videos, the typical gazer head
186 velocity before the head stopped. This was accomplished by lining up all videos based on the
187 head stop and averaging the head velocities (Figure 3c).

188



(a)

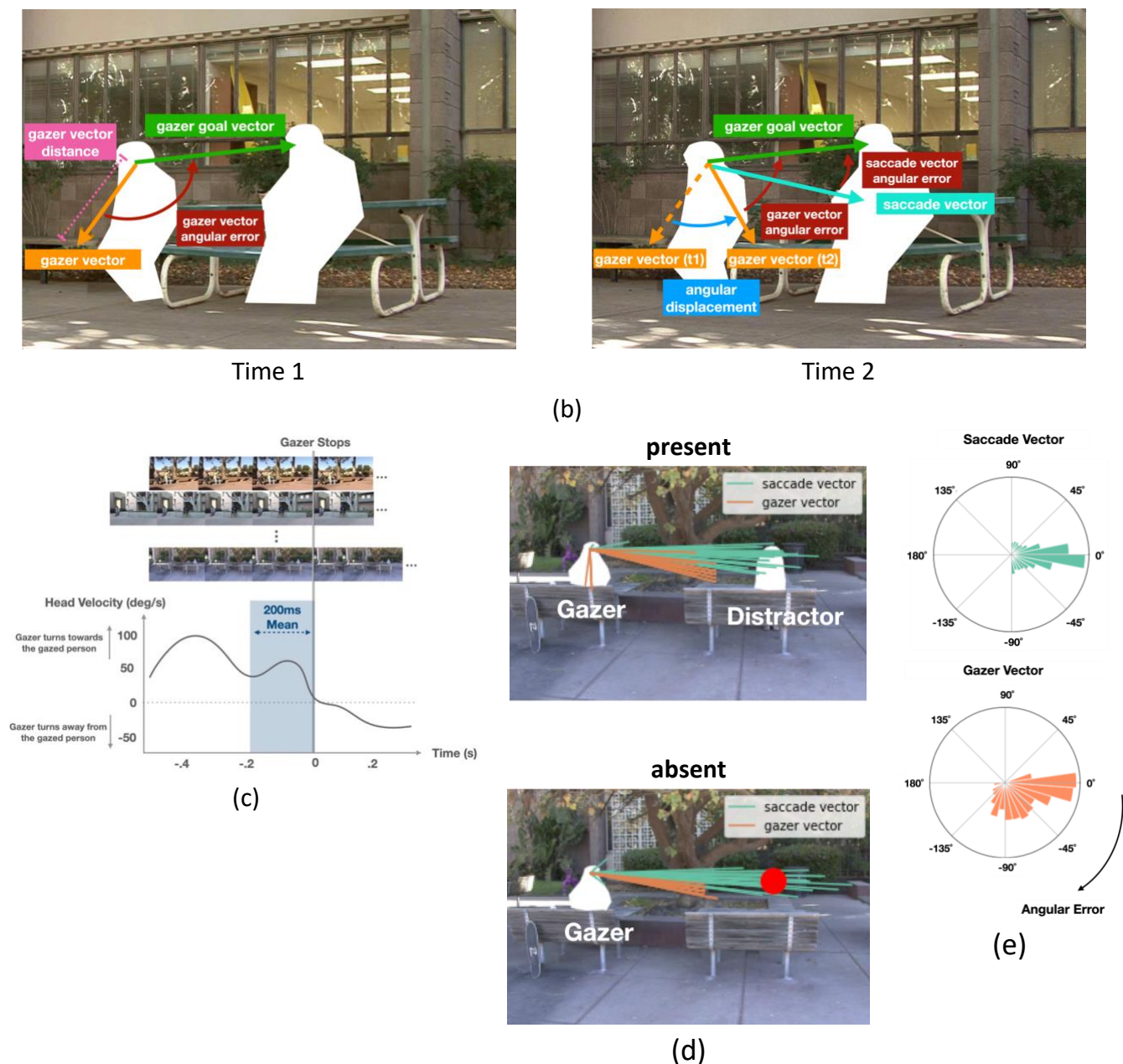


Figure 3. (a). Workflow for AI model for gaze estimation (Chong et al., 2020). The model takes individual frames, paired with a binary mask that indicates the gazer’s head location within the scene, and a cropped image of the gazer’s head, to produce a probability heatmap. The pixel location with the highest probability was taken as the final estimated gazed location and gazer vector endpoint (orange arrow in final estimation image). We computed various frame-to-frame gaze features based on the gazer vectors and related them to the dynamics of observers’ eye movements during gaze-following. (b). Examples of the initial *gazer vector*, the *gazer vector distance*, the *gazer goal vector*, the *angular displacement*, and *angular errors*. The *gazer vector distance* was the vector length indicating how far away the estimated gazed location was from the gazer. The *gazer goal vector* is the vector whose start point was the gazer’s head centroid and the endpoint was the gazer goal location. The *angular displacement* is the angle between the current *gazer vector* and the initial *gazer vector*

position. The *angular error* is the angle between the current *gazer/ saccade vector* and the *gazer goal vector*. c) Estimation of the typical head velocities right before (200ms interval) the gazer's head stops moving. Velocities were obtained by aligning all videos with respect to the gaze stop time and averaging the head velocities. Head velocity = 0 at time = 0. (d). The first saccade vectors (teal lines) and corresponding gazer vectors (orange lines) at the saccade initiation times for all observers and trials for the same video (top: present condition, bottom: absent condition). (e). Histogram of angular errors for first saccade vectors and gazer vectors at the saccade initiation times for all trials. All vectors were registered relative to the gazer goal vector (the horizontal direction to the right represents 0 angular error).

189

190 ***Anticipatory first saccades that predict gaze goal direction***

191 The gazers' head movements started with the video onset and their mean duration was 0.61s.
192 The observers' mean first saccade latency was 0.34s (std=0.07s). Thus, the saccade initiation
193 most often preceded the end of the gazer's head movement. In 81% of the trials, participants
194 initiated the first saccade before the gazer's head movement stopped (86% of the trials for target-
195 present, 85% for distractor present, and 77% for target/distractor absent trials). We investigated
196 whether these anticipatory first saccades were based on a prediction beyond the available
197 information in the gazer's head direction at the time of saccade initiation. Or on the contrary, are
198 the saccade directions based on the information in the gazer's head direction at the time of
199 saccade initiation?

200

201 To evaluate these hypotheses, we measured the angular error between the DNN-estimated
202 gazer's head direction (gazer vector) at the time of the first saccade initiation and the gazer goal
203 vector (Figure 3b right) for each trial. The gazer vector angular error at the time of saccade
204 initiation provides a lower bound on observers' saccade angular error if the brain only used the
205 gazer's head direction to program the eye movements. Figure 3d visualizes the first saccade
206 vectors (teal lines) and corresponding gazer vectors (orange lines) at the saccade initiation times
207 for all observers and trials for a sample video. The results show how the saccade directions are
208 closer to the gazer goal direction than the direction information provided by the gazer's head at
209 the time of saccade initiation (gazer vector). Figure 3e shows co-registered saccade vectors and
210 gazer vectors at the time of saccade initiation across all trials/observers. The horizontal line
211 pointing to the right represents zero angular error (i.e., a saccade or gazer vector that points in
212 the same direction as the direction of the gazer goal). The mean angular error for the saccade
213 directions was significantly smaller than that of the gazer vector at the time of saccade initiation
214 (18 degrees vs. 40 degrees, bootstrap $p < 1e-5$). This difference was larger for target/distractor
215 present videos (14 degrees vs. 42 degrees, bootstrap $p < 1e-5$) but was still significant even when
216 the target/distractor was absent (22 degrees vs. 38 degrees, bootstrap $p < 1e-5$). The findings
217 suggest that observers make anticipatory first saccades that infer the direction of the gaze goal
218 beyond the momentary information from the gazer's head direction. We estimated the
219 additional time after saccade initiation it took for the gazer's head to point in the direction of the
220 saccade. On average it took 0.37s (std across observers=0.11s) and 0.22s (std across
221 observers=0.09s) for the gaze vector to reach the saccade vector direction for videos with
222 target/distractor present and target/distractor absent respectively.

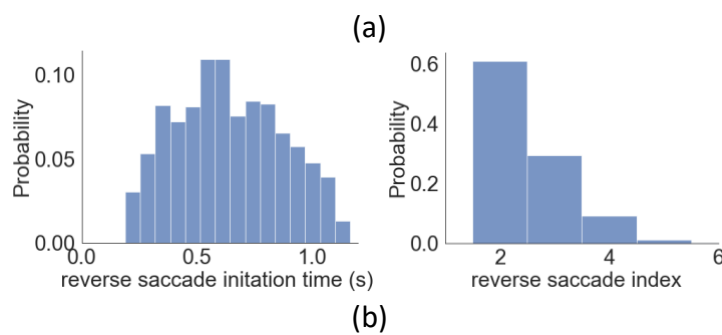
223
224 To make sure the results were not due to inadequate gaze estimates by the DNN, we repeated
225 the analysis with humans-estimated gazer vectors instead of DNN-estimated gazer vectors. The
226 human-estimated gazer vectors were obtained from ten individuals (not participants in the study)
227 that viewed randomly sampled individual frames from the videos and were instructed to select
228 the gaze goal (see methods). Because we were interested in measuring the inherent information
229 provided by the gazer's head direction independent of the peripheral information, the
230 participants viewed frames from the target/distractor absent videos. The human-estimated gazer
231 vectors resulted in smaller angular errors than the DNN but showed similar findings. Observers'
232 mean first saccade angular error was significantly smaller than the mean human gazer vector
233 angular error (18 degrees vs. 32 degrees, bootstrap $p < 1e-5$). This effect was present for both, the
234 target/distractor present videos (14 degrees vs. 36 degrees, bootstrap $p < 1e-5$) as well as the
235 target/distractor was absent (22 degrees vs. 27 degrees, bootstrap $p = 0.017$). On average it took
236 0.34s (std=0.12s) and 0.16s (std=0.09s) and for the gazer vector to reach the 1st saccade vector
237 location for the present and absent conditions.
238

239 ***Frequent reverse saccades triggered by gazer's low head velocity***

240 Even if we explicitly instructed participants to follow the gaze, our analysis of eye position
241 revealed that participants executed backward saccades in the opposite direction of the gazer
242 vector (reverse saccades) in 22% of all trials (see Figure 4a for an example). The mean reverse
243 saccade initiation time was 0.63s (std= 0.07s, Figure 4b). Over 80% of the reverse saccades were
244 either the second or the third saccade in the trial (reverse saccade index, Figure 4b). The mean
245 duration of the gazer's head movement during reverse saccade trials was 0.65s. In 87% of the
246 videos, the gazer started to look away from the gazer person at the end of the movie (DNN
247 estimation mean=0.98s, std=0.18s, human estimation mean=1.06s, std=0.15s). In those videos,
248 the majority of reverse saccades (88%) were executed before the gazer started looking away.
249 Figure 4c shows the frequency of first saccades and reverse saccades, as well as the overall head
250 velocity over time. Trials with reverse saccade had significantly shorter first saccade latencies
251 compared to those without reverse saccade (Figure 4d, target/distractor present condition 0.23s
252 vs. 0.34s, $p = 2.6e-67$, absent condition 0.27s vs. 0.40s, $p = 4.2e-50$, both posthoc pairwise
253 comparison with FDR). What could explain the shorter first saccade latencies of trials with reverse
254 saccades? One possible interpretation is that early first saccades are unrelated to the stimulus
255 properties and are generated by stochastic processes internal to the observer. Consequently,
256 when the first saccade is executed too early, a compensatory reverse saccade is subsequently
257 programmed.
258

259 An alternative possibility is that the observer's early first saccade executions are not random but
260 related to some aspect of the gazer's head movement. To investigate this possibility, we first
261 analyzed the average head velocity over time relative to the timings of the video onset
262 (coincident with the gazer head movement onset) and first saccade execution. The analysis was

263 done separately for trials with and without reverse saccade. If the early first saccades in reverse
264 saccade trials are triggered randomly and are unrelated to the gazer's head features, we should
265 find no significant difference in average head velocity between the two types of trials. Instead,
266 we found a significantly lower head velocity during the first 0.23s of the video for the trials with
267 reverse saccades, 63.6 deg/s vs. 93.6 deg/s, cluster-based permutation test, $p=1.0e-04$ (Figure
268 4e, average head velocity lined up with video start). When we aligned the data with the initiation
269 time of the first saccade, we also observed a significantly lower head velocity for the trials with
270 reverse saccade during 0.37s before the first saccade initiation 47.1deg/s vs. 104.9deg/s, cluster-
271 based permutation, $p=1.0e-4$ (Figure 4f). Furthermore, the average head velocity of 47.1 deg/s
272 was within the range [31.6 deg/s-62.5 deg/s, 95 % confidence interval] of the average head
273 velocity before the gazer's head stops (estimated from all movies; see the horizontal green band
274 in Figure 4e,f). These findings suggest that when the gazer's head velocity is slow, observers make
275 an inference that the gazer might be stopping their head movement. Observers then execute an
276 eye movement to the currently estimated gazer goal. Thus observers' faster first saccades are
277 not executed at random times but are related to the observers' inference that the gazer's head
278 movements might come to a stop. Figure 4f also shows that right before the execution of the first
279 saccade, in reverse saccade trials, the head velocity starts accelerating. We interpret this to
280 indicate that observers infer from the accelerating velocity just prior to the first saccade
281 execution that the gazer's head will not come to a stop. Consequentially, observers program a
282 reverse saccade.
283



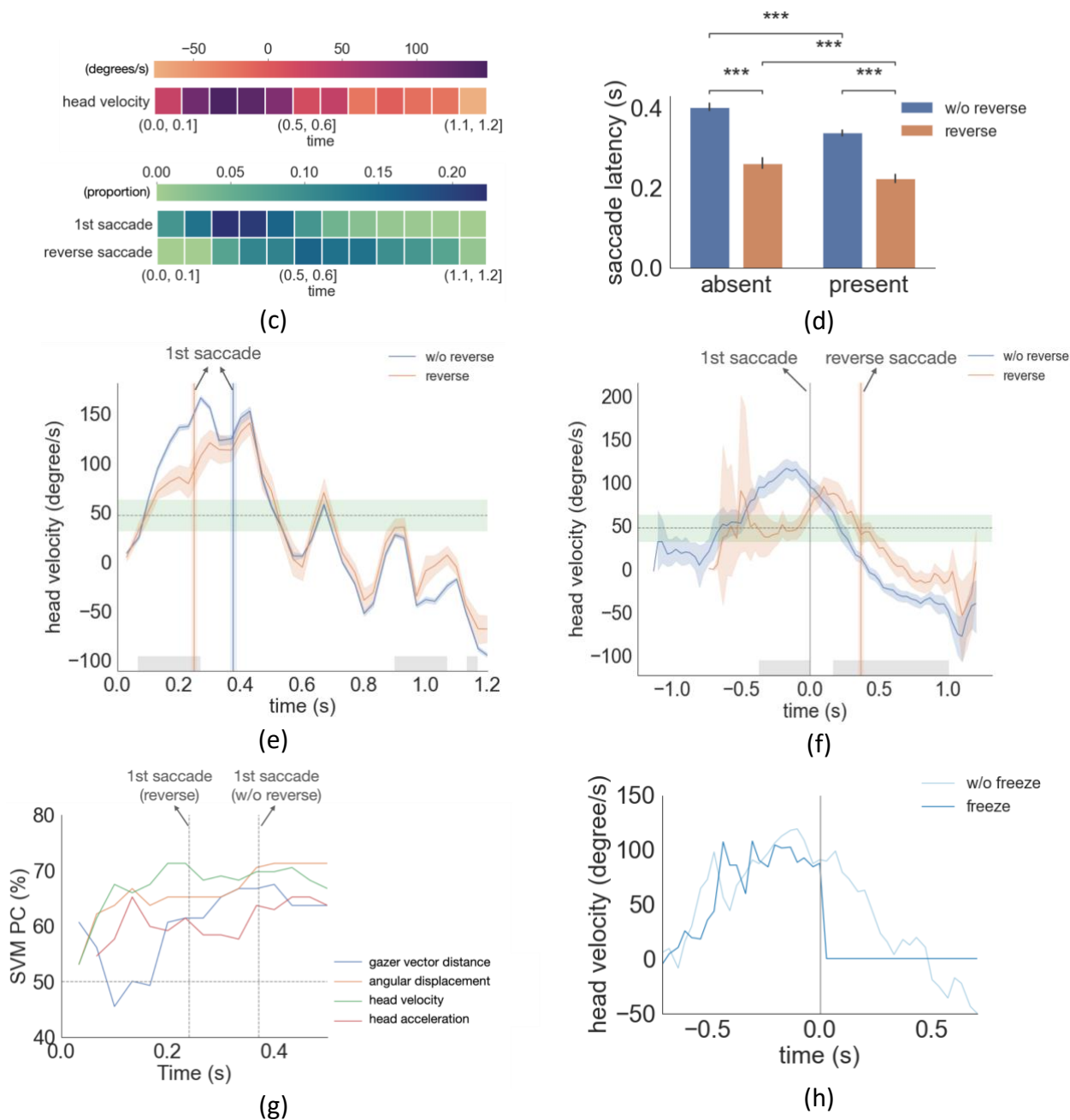


Figure 4. (a) An example of eye movement trace for three saccades over time. A first gaze-following saccade, followed by a reverse saccade, and another post-reverse saccade gaze-following saccade. The light red to dark red represents the order in time (see video demo at <https://osf.io/yd2nc>). (b) Histogram of the reverse saccade initiation time and reverse saccade index (2nd, 3rd, etc.). (c) Heatmaps represent the first saccade and reverse saccade frequency, and the gazer's head velocity over time (d) Saccade latency separated by three conditions and reverse saccade trials. (e) Gazer's head rotation velocity vs. time separated for reverse saccade and non-reverse saccade trials. Shaded areas are the 95% bootstrapped confidence interval. Positive velocity represents moving toward the gazed person's location. The vertical lines are the mean first saccade latency. The gray area shows the statistical significance under the

cluster-based permutation test. The green area represents the 95% confidence interval of the velocity right before the gazer's head stops moving across all movies. (f) The same figure as (e) except that head velocity was aligned at the initiation time of the first saccade. (g) The proportion correct of linear SVM models trained to predict whether a movie was in the upper 50 %/or lower 50 % of movies with reverse saccades. The x-axis is the time range from the movie used to train the SVM model. The first saccade latency and reverse saccade latencies are marked as dashed lines as references. (h) The head velocity aligned with the first saccade initiation time at $t=0$, separately for trials with frozen frames and without.

284
285 Our analysis focused on the head velocity, but what about other features of the gazer's head?
286 The supplementary material S1 shows analyses for other features including distance, angular
287 displacement, and head acceleration. Other features are also significantly different across
288 reverse and non-reverse saccade trials. This is not surprising because there is a correlation
289 between some of the features. For example, before the first saccade execution, the angular
290 displacement is smaller for reverse saccade trials. This is because slower angular velocity for the
291 head will result in lower angular displacement at the time before the first saccade. Still, head
292 velocity showed the clearest results. In order to further investigate whether the head velocity or
293 other gaze features from the videos can better explain reverse saccades, we trained multiple
294 support vector machine (SVM) models using different head features to predict the frequency of
295 reverse saccades (binary prediction: top vs. bottom 50 percentile) using features: 1. Gazer vector
296 distance 2. Angular displacement, 3. Head angular velocity, 4. Head angular acceleration (see
297 Methods for detailed description). We used the time range starting from the beginning of the
298 video and gradually increased the time range for the predictor, and plotted the SVM model
299 proportion correct (PC) in Figure 4g. We found that the head velocity had the highest accuracy in
300 predicting reverse saccade movies among all gaze features. The model's accuracy peaked when
301 we used head velocity information from 0-230ms of each video (71.2%) and asymptoted
302 afterward. This was consistent with the results that during the first 0.23s of movies, trials with
303 reverse saccade had a significantly lower head velocity than those without.

304 305 ***When are the reverse saccades planned?***

306 Having established that the gazer's low head velocity might be triggering an early first saccade in
307 trials with reverse saccades, we tried to determine the timing of the reverse saccade
308 programming. One possibility is that the reverse saccades are programmed after the execution
309 of the first forward saccade. In this framework, the gazer's initial slow head velocities in some
310 trials trigger an early first saccade forward and, during that subsequent fixation, the motion of
311 the gazer's head accelerating captures attention and triggers the reverse saccade. A second
312 possibility is that it is the gazer's head velocity increase right before the observer executes the
313 first saccade (Figure 4f) that triggers the programming of the reverse saccade prior to the

314 execution of the first forward saccade. To assess these two hypotheses, we conducted another
315 experiment, in which we monitored in real time the eye position of observers and froze the video
316 frames immediately after participants initiated the first gaze-following saccade (Figure 4h, See
317 demo video at <https://osf.io/xpueh>). This only occurred randomly in 50 % of the trials to prevent
318 observers from changing their eye movement strategy. If observers' reverse saccades were
319 triggered by the transient motion after the first saccade execution, then freezing the video and
320 eliminating the transient peripheral motion signal of the head should diminish the frequency of
321 the reverse saccades. However, we found that freezing the video frame after the first saccade
322 execution did not reduce the proportion of trials with reverse saccade relative to the unfrozen
323 videos trials, (mean=22%, std=12% for frozen vs. mean=21%, std=11% for unfrozen, bootstrap
324 $p=0.6$). These results suggest that observers planned the reverse saccade prior to the execution
325 of the first forward saccade.

326

327 ***Functional role of reverse saccade***

328 Next, we tried to understand the function, if any, of reverse saccades. We first analyzed the
329 endpoint of the reverse saccade. We found that the reverse saccades landed close to the gazer's
330 head (Figure 5a; mean distance to the gazer's head 0.79° , std= 0.28°) suggesting that the reverse
331 saccades aim to re-fixate the gazer given the change in the gazer vector after execution of the
332 first saccade. To assess the potential functionality of the reverse saccade, we compared the error
333 in fixating the gaze goal (saccade error: saccade endpoint distance to the gazed person's location)
334 of forward saccades before and after the reverse saccade. Figure 5b shows the density map of
335 forward saccade endpoints separately for pre and post reverse saccades for a single sample
336 image, as well as the density map combined across all images by registering the saccade
337 endpoints relative to the gazer's head. Forward saccades following a reverse saccade ended
338 closer to the gaze goal than the saccades before reverse saccades. (Figure 5c; 1.8° for post reverse
339 saccade vs. 2.5° for pre reverse, $p=0.0054$ based on bootstrap resampling, see methods). For the
340 target/distractor absent condition, we did not find this effect, 5.2° vs. 5.3° , $p=0.43$. Finally, the
341 saccade error in the trials without reverse saccades was significantly lower compared to the trials
342 with reverse saccades (present 1.1° , absent 3.7° , all $p<0.001$, corrected by FDR). This suggests
343 that the gazer information was less ambiguous and more accessible to observers in the trials with
344 no reverse saccades.

345

346

347

348

349

350

351

352

353

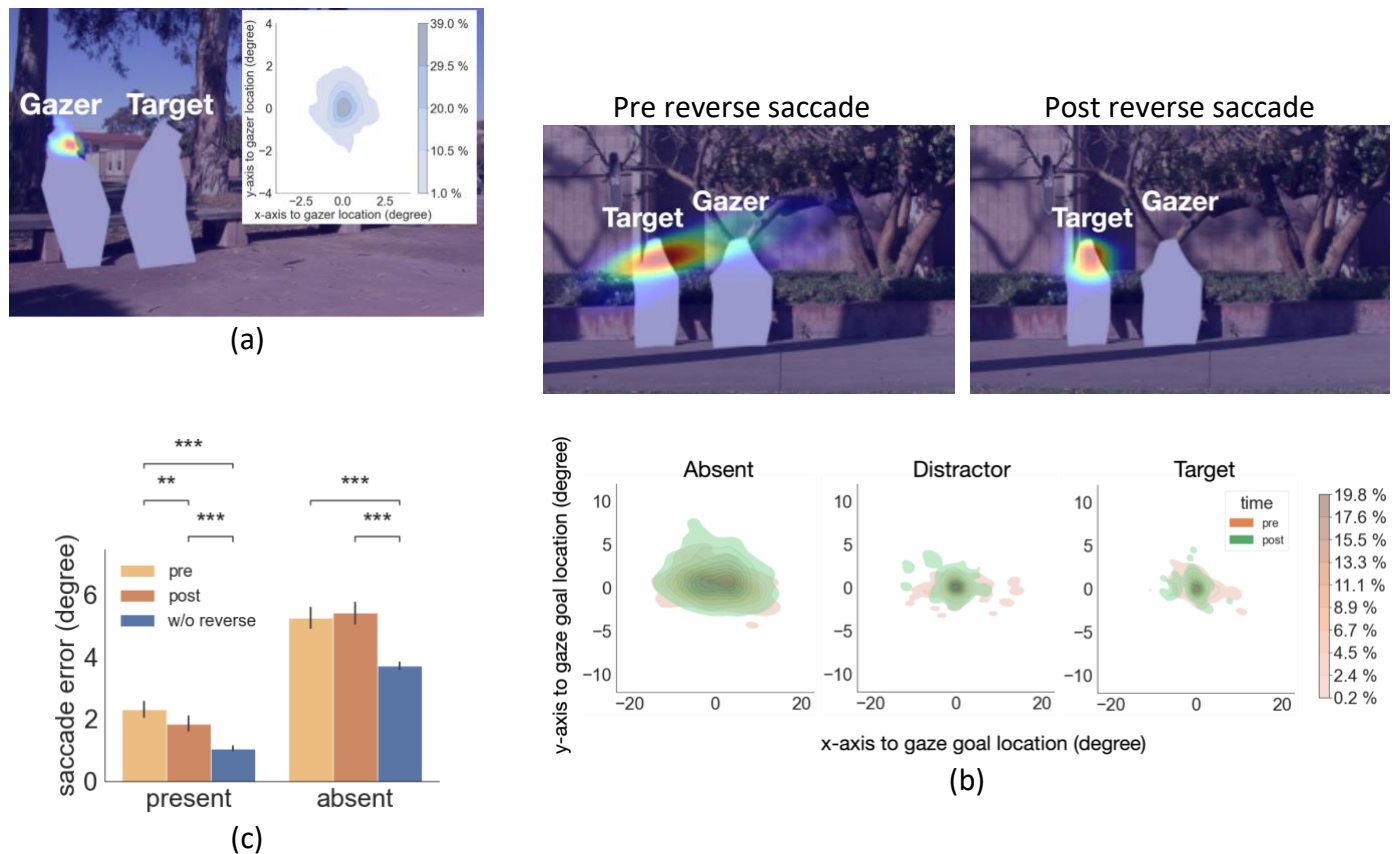


Figure 5. (a) Density map of reverse saccade endpoint locations overlaid an example image. The density map of all reverse saccade locations registered across videos relative to the gazer's head location at origin (0,0) is shown on the top right. Colorbar shows the proportion of saccades falling in each region. (b) Top: Density map of gaze-following saccade location pre- and post-reverse saccade overlaying on an example image. Bottom: Density map of all saccades pre- and post-reverse saccades registered relative to the gazed person's head location at origin (0,0). Colorbar shows the proportion of saccades falling in each region. (c) The saccade error (relative to the gazed location, center of the head) for pre- and post-reverse saccades. Trials with no reverse saccade were treated as the baseline condition.

354

355

356 ***Causal influence of re-foveating the gazer with a reverse saccade***

357 Our analysis showed that the saccade endpoint after the reverse saccade was closer to the gaze
 358 goal than the endpoint of the forward saccade preceding the reverse saccade. The interpretation
 359 is that re-fixating the gazer with the reverse saccade improved the inference about the gazer goal
 360 and benefited the subsequent forward saccade. However, an alternative explanation is that the
 361 gaze-following saccade after a reverse saccade simply has longer visual processing compared to
 362 the first saccades preceding the reverse saccades (first saccade initiation time $m=0.35s$ vs. first

363 saccade post reverse saccade initiation time $m=0.84s$). Longer processing times would result in
364 better estimates of the gaze goal.

365
366 To assess these two competing explanations for the reduction of error of gaze-following saccades
367 after a reverse saccade, we implemented a follow-up experiment with twenty-five new observers.
368 In the new experiment, we digitally erased the gazer on 50 % of the reverse saccade trials
369 (randomly) before the re-fixation of the gazer. To accomplish this, we monitored eye position in
370 real-time, and whenever we detected a reverse saccade during the video, we erased the gazer
371 with a 50 % probability. The experiment allowed comparing the errors of gaze-following saccades
372 subsequent to a reverse saccade with matched visual processing times.

373
374 If the reduced saccade error is related to the foveal re-processing of the gazer after the reverse
375 saccade, we should expect a larger saccade error when we erase the gazer (see Figure 6a
376 example). We first confirmed that the basic analyses replicated the first experiment. The mean
377 reverse saccade initiation time was 0.69s (std = 0.07s), with 80 % of the reverse saccades being
378 the second or the third saccade. Reverse saccades occurred in 31 % of the trials. Trials with
379 reverse saccade had a significantly smaller first saccade latency compared to those without
380 reverse saccade (0.23s vs. 0.33s, bootstrap $p<1e-5$). Reverse saccade trials were associated with
381 slower head velocity during the initial period of the movie (100ms-260ms) and 150 ms before the
382 first saccade (Supplementary material S2). For trials without the gazer removed, we found the
383 same effect of smaller saccade error post reverse saccade for both the present condition (1.4° vs.
384 1.8° , bootstrap $p=1.2e-4$) and the absent condition (4.8° vs. 5.2° , bootstrap $p=0.04$), and (Figure
385 6b).

386
387 Critical to our hypotheses testing, the results showed that the saccade error post-reverse
388 saccade was significantly higher in the trials with the gazer removed compared to those with
389 unaltered videos, for both the target/distractor present condition (2.5° vs. 1.4° , bootstrap $p<1e-5$)
390 and the absent condition (5.1° vs. 4.8° , bootstrap $p=0.006$; Figure 6c). The time of the forward
391 saccade following the reverse saccade was the same across trials with the gazer removed or
392 unaltered (0.83s from video onset with the gazer unaltered vs. 0.8s, with the gazer removed,
393 bootstrap $p=0.1$). This finding confirms that the benefit of reducing the gaze-following saccade
394 errors is causally linked to the uptake of additional gaze goal information from re-fixating the
395 gazer through a reverse saccade.

396
397
398
399
400

401

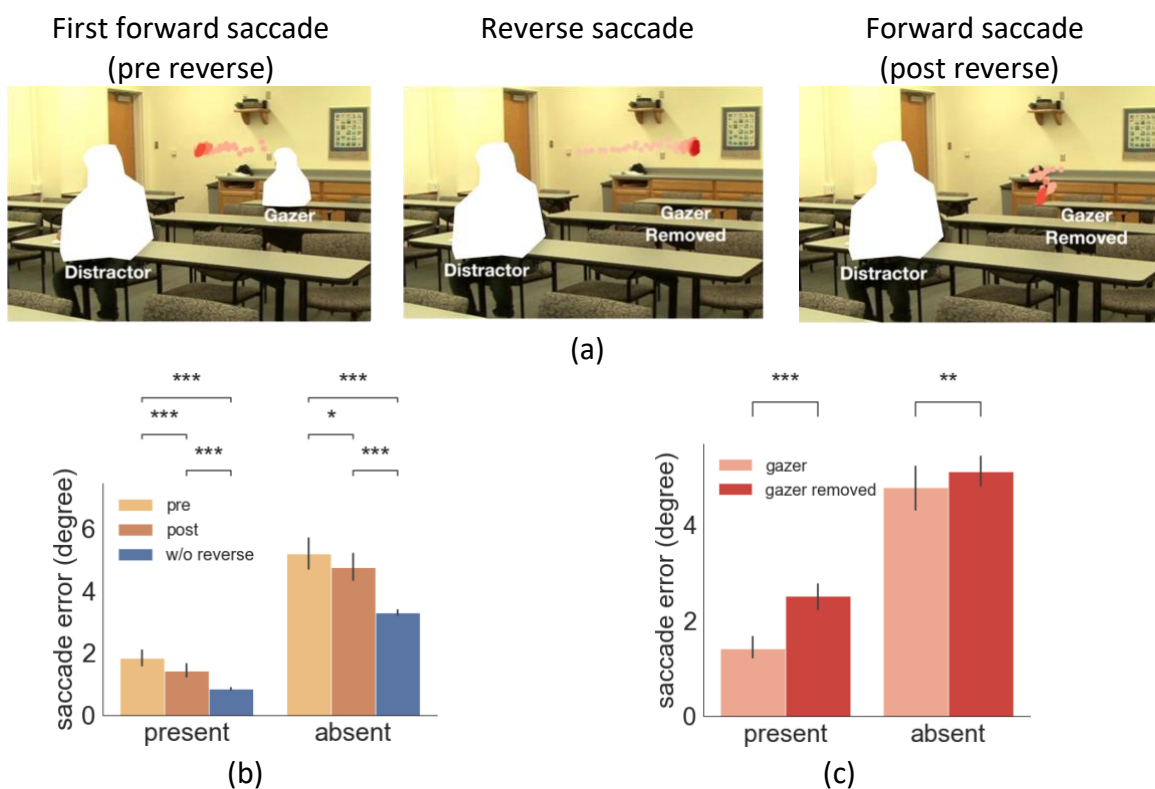


Figure 6 (a) Example of eye movement trace over time when the gazer was erased triggered by a reverse saccade. The light red to dark red represents the order in time (See video demo at <https://osf.io/etqbw>). (b) The saccade error (relative to gazed person's head) in trials without gazer removed pre reverse saccade vs. post reverse saccade vs. baseline trials (w/o reverse saccades). (c) The saccade error post reverse saccade with gazer removed vs. gazer unaltered.

Anticipatory and reverse saccades during free-viewing search

402 Our two experiments instructed observers to follow the gaze of the person in the video. This
 403 might be unnatural, and might have motivated observers to follow the gazer's head movements
 404 and trigger anticipatory saccades and reverse saccades. To assess the generality of our findings
 405 we implemented a control experiment (Experiment 3) with five participants where we did not
 406 explicitly instruct observers to follow the gaze during the video presentation. Instead, we only
 407 instructed them to evaluate whether they could find the target person and decide whether they
 408 were present. No information was given about the gazer or eye movement strategies to follow.
 409 We found that participants spontaneously executed gaze-following saccades for 74% (std = 10%)
 410 and 91% (std=7.4%) of the trials for the absent and the present condition, respectively (S3).
 411 Observers also executed anticipatory first saccades prior to the end of the gazer's head
 412 movement on 88% of the trials. We also observed an even larger number of reverse saccades

413 than in the first experiment where observers were instructed to follow the gazer (33%, std=22%
414 and 37%, std=16%) of trials for the absent and the target or distractor present condition,
415 respectively). These findings suggest that anticipatory and reverse saccades are not a byproduct
416 of the instructions in experiment 1.

417

418 **Discussion**

419 We investigated eye movement control while following the gaze of others. Although human eye
420 movements are fast and might seem idiosyncratic, our findings show that the human brain uses
421 moment-to-moment information about the gazer's head dynamics and peripheral information
422 about likely gaze goals to rationally plan the timing and endpoint of saccadic eye movements.
423 First, we found that the oculomotor system integrates information about the foveally presented
424 gazer's head and peripheral information about potential gaze goals. When a person was present
425 at the gaze goal, observers executed faster and more accurate saccades. The findings are
426 consistent with a series of studies showing observers' ability to simultaneously process foveal
427 and peripheral information for simpler dual tasks with simple stimuli (Ludwig et al., 2014; Stewart
428 et al., 2020) and their joint influence on fixation duration during scene viewing (Laubrock et al.,
429 2013) and subsequent eye movements (Wolf et al., 2022). Importantly, the first saccades are
430 anticipatorily initiated before the gazer's head movement comes to a stop. And they contain
431 information about the direction of the gaze goal that is more accurate than the direction
432 information provided by the gazer's head at the time of saccade initiation. This suggests that the
433 brain is using peripheral information to make an active prediction about likely gaze goals.
434 Furthermore, previous studies have shown that a saccade is typically based on visual information
435 presented ~ 80 ms prior to saccade execution (Becker & Jürgens, 1979; Caspi et al., 2004; Hooge
436 & Erkelens, 1999; Ludwig et al., 2005). Thus the first saccade might only have access to the gazer's
437 head direction up to ~80 ms before saccade execution. This inferential process when a person is
438 present in the scene might be expected. But, the finding still prevailed when a target/distractor
439 was absent. It is likely that even when no person is present at the gaze goal location, the brain
440 uses information about the scene including the ground, the objects, and the sky to make
441 estimates of likely gaze goals. Prior knowledge about the maximum angular rotation of the
442 gazer's head also constrains the possible gazer goals.

443

444 Second, we found that early first saccades are executed when the gazer's head velocity
445 diminishes to values comparable to the velocity that is typical during the 200 ms time interval
446 before the head stops. This is consistent with the idea that observers are using the gazer's head
447 velocity to dynamically make inferences about the likelihood that the head will stop. However,
448 our data also suggest that other cues are used to infer that the gazer's head will stop. For example,
449 for some trials with longer first saccade latencies (no reverse saccade trials), the head velocity

450 before the saccade execution is almost double the typical head velocities during the 200 ms time
451 interval before a head stops (Figure 4e). Thus, the observers must rely on other cues. In these
452 long latency trials (Figure 4e) there is a reduction of the head velocity in the 200 ms before the
453 saccade execution suggesting that observers use the head's deceleration to infer that the gazer's
454 head will come to a stop and then execute the first gaze-following saccade. It is also likely that
455 for trials with a gaze goal, observers use an estimated error between the implied gaze direction
456 and the gaze-goal to plan saccades. Small estimated angular errors might be used to trigger
457 saccades. Thus, we suggest that the oculomotor system might use multiple cues (head velocity,
458 head deceleration, estimated gaze errors, etc) to trigger gaze-following saccades.

459
460 Third, surprisingly, we found that observers often executed reverse saccades in a significant
461 proportion of trials (> 20 %). The reverse saccades were not an artifact of our instruction to the
462 observers to follow the gaze of the person in the video. A follow-up experiment where observers
463 were instructed to decide whether a target person was present with no instruction about eye
464 movements also resulted in a comparable proportion of reverse saccades. Why might observers
465 make such saccades? Our analysis showed that these reverse saccades do not appear randomly
466 across trials. Reverse saccades occur on trials in which the gazer's head velocity is slow but starts
467 accelerating about 200 ms before the first saccade is executed and observers infer that the
468 gazer's head will not come to a halt. Why don't observers simply cancel the forward saccade?
469 Studies have shown that there is an 80-120 ms delay between the programming of a saccade and
470 its execution (Becker & Jürgens, 1979; Caspi et al., 2004; Ludwig et al., 2005). The gazer's head
471 acceleration occurring immediately before the execution of the forward saccade is not used to
472 cancel the impending planned eye movement.

473
474 Our findings with the experiment that freezes the gazer after the first forward saccade suggest
475 that the reverse saccade is programmed prior to the execution of the first forward saccade. This
476 concurrent programming of saccades has been documented for simplified lab experiments
477 (Becker & Jürgens, 1979; Caspi et al., 2004) but not in the context of real-world stimuli and tasks.
478 One alternative explanation we did not explore is that reverse saccades are simply triggered after
479 first forward saccades that do not land on the target/distractor. In this perspective, a forward
480 saccade is executed and when foveal processing determines that the saccade endpoint was far
481 from a likely gaze goal then a reverse saccade is programmed and executed (regardless of the
482 velocity of the gazer's head). Data analysis does not favor this interpretation. In a small
483 percentage of trials (15 % of reverse saccade trials) first saccades landed within 0.5° visual angle
484 of the target but these were still followed by reverse saccades. This observation suggests the
485 presence of the close-to-fovea gaze goal was not sufficient to interrupt a reverse saccade
486 programmed prior to the execution of the first forward saccade.

487

488 Our results also show that the reverse saccades had functional importance as forward saccades
489 after re-fixating the gazer were more accurate at landing close to the gaze goal. The benefit of
490 re-fixating the gazer was more reliable when there was a person present at the gaze goal. When
491 a gaze goal person was absent we found a less reliable re-fixation benefit (not statistically
492 significant in experiment 1 and marginally significant in experiment 2) suggesting that not having
493 a peripheral likely gaze goal can be a bottleneck to the accuracy of saccade endpoints.

494
495 The existence of reverse saccades might seem puzzling. Why does the oculomotor programming
496 system not wait longer until the gazer's head comes to a full stop, then executes the gaze-
497 following saccade and avoids programming reverse saccades altogether? Executing anticipatory
498 eye movements that predict future grasping actions (Mennie et al., 2007; Pelz & Canosa, 2001),
499 the location or motion of a stimulus (Fooker & Spering, 2020; Kowler, 1989, 2011; Kowler et al.,
500 2019) is common for the oculomotor system. Thus, while following the gaze of others the
501 oculomotor system plans anticipatory saccades before the completion of the gazer's head
502 movements that predict the gaze goal. Occasionally, these predictive saccades are premature
503 and the brain rapidly programs a reverse saccade to refixate the gazer and collect further
504 information about the potential gazer goal.

505
506 Are the reverse saccades unique to gaze following? No, humans make reverse saccades in other
507 visual tasks that require maintaining information in working memory, such as copying a color
508 block pattern across two locations (M. Hayhoe & Ballard, 2005; M. M. Hayhoe, 2017;
509 Meghanathan et al., 2019). Most notably during reading humans make frequent reverse saccades
510 ("called regressive"). Although one might draw a parallel between reading and gaze-following,
511 our findings highlight important distinctions. Regressive saccades during reading are related to
512 inaccurate eye movements that missed critical words or fixations that are too short to deeply
513 process a word's meaning (Inhoff et al., 2019; Rayner, 1998). The reverse saccades while
514 following dynamic gaze are related to moment-to-moment changes in the visual information in
515 the world (i.e. the gazer's head velocity) and the oculomotor systems' rapid strategy changes to
516 optimize gaze-following.

517
518 What might be the brain areas involved in the oculomotor programs for gaze following? There is
519 a large literature relating gaze position to neuronal response properties in the superior temporal
520 sulcus (Oram & Perrett, 1994) and dorsal prefrontal cortex (Lanzilotto et al., 2017). These areas
521 relay information to the attention and gaze network in the parietal and frontal cortex which are
522 responsible for covert attention and eye movements (Pierrot-Deseilligny et al., 2004). Finally, the
523 concurrent programming of saccades has been related to neurons in the Frontal Eye Fields (FEF,
524 (Basu & Murthy, 2020)). Identifying brain areas that integrate peripheral information to generate
525 predictions of likely gaze goals is an important future goal of research.

526
527 One possible limitation of our study is that it focused on the head movement while a large
528 literature focuses on the influences of the gazer's eyes (Driver et al., 1999; Friesen et al., 2004;
529 Langton et al., 2000; Mansfield et al., 2003; Ristic et al., 2002). Our study was relevant to gazers
530 situated at a distance from the observers. The mean angle subtended by the heads in our
531 videos (1.47° , $\text{std}= 0.32^\circ$) would match the angle subtended by a real-sized head viewed at a
532 distance of 9.3 m ($\text{std}= 2.0$ m) in real life. At that distance, the eye subtends a mean angle of
533 0.147° (vertically) providing a poor source of information to infer the gaze goals compared to
534 the head orientation. Future studies should investigate gazers at smaller distances from the
535 observer and assess how dynamic gazer eye and head movements are integrated and their
536 interactions (for static images see Balsdon & Clifford, 2018; Cline, 1967; Langton, 2000; Langton
537 et al., 2004; Otsuka, 2014). Similarly, we did not analyze lower body movements. Recent studies
538 have shown the diminished influence of the lower body on the orienting of attention (Han &
539 Eckstein, 2022; Pi et al., 2020).

540
541 To conclude, our findings reveal the fine-grained dynamics of eye movements while following
542 gaze and the inferential processes the brain uses to predict gaze goals and rapidly program
543 saccades. Given that attending to the gaze of others is an integral part of a normal functioning
544 social attention system, our findings might provide new granular analyses of eye movement
545 control to assess groups with social attention deficits for which simpler gaze following analyses
546 have shown disparate results (Chawarska et al., 2003; Nation & Penny, 2008; Ristic et al., 2005).

547

548 **Materials and Methods**

549 *Experiment 1*

550 **Subjects**

551 Experiment protocols were approved by the University of California Internal Review Board.
552 Twenty-five undergraduate students (ages 18-20, 16 females, 9 males) from the University of
553 California Santa Barbara were recruited as subjects for credits in this experiment. All have
554 normal to corrected-to-normal vision. All participants signed consent forms to participate in the
555 study.

556

557 **Experimental setup and stimuli**

558 All videos were presented at the center of a Barco MDRC 1119 monitor with 1280×1024
559 resolution, subtending a visual angle of $18.4^\circ \times 13.8^\circ$ (width x height). Participants' eyes were
560 75cm from the computer screen with the head positioned on a chin rest while watching the
561 videos (0.023° visual angle/pixel). Each subject's left eye was tracked by a video-based eye
562 tracker (SR Research Eyelink 1000 plus Desktop Mount) with a sampling rate of 1000Hz.

563 Subjects' eye movements were calibrated and validated before the experiment. Any large eye
564 drifts that caused failure in maintaining fixation at the beginning of each trial ($> 1.5^\circ$ visual
565 angle) would result in observers having to do a recalibration and revalidation. Events in which
566 velocity was higher than $35^\circ/\text{s}$ and acceleration exceeded $9500^\circ/\text{s}^2$ were recorded as saccades.

567
568 Stimuli consisted of 80 videos (1.2s long) originally taken from videos recorded at the University
569 of California Santa Barbara campus in different settings (classrooms, campus outdoors, student
570 apartments, etc.). During the filming, we gave verbal instructions to the actor to look toward
571 another person. Once the video starts, one gazer initiated the gazing behavior (looking at
572 another person) toward either a distractor person (50% chance) or a target person (50%
573 chance). The mean eccentricity of the gazed person relative to the gazer was 6° , $\text{std}=3^\circ$, with a
574 minimum of 1.3° and a maximum of 13.6° (Supplementary material S4). The gazed person was
575 either present in the video (original) or was erased from the video and appeared invisible.
576 Therefore, in total there were 80 videos x 2 (present vs. absent) = 160 video stimuli. To erase
577 the gazed individuals from the images, we replaced the RGB values of pixels contained by the
578 individual outline (annotated by research assistants) with the RGB values of those pixels of the
579 immediate background (Figure 1a-b). The gazed person's location relative to the gazer's head
580 had a mean of 6° visual angle, $\text{std}=3^\circ$ visual angle (Figure 1c).

581
582 Across all the movies, the head regions subtended a mean size of 1.47° ($\text{std}=0.32^\circ$). Given that
583 the average vertical length of eyes spans 2.4cm (0.024m) (Bekerman et al., 2014) and the
584 average vertical distance of the head is about 0.24m (Lee et al., 2006), the eye only
585 spanned $.147^\circ$ ($\text{std}=0.032^\circ$).

586 587 Procedure

588 Subjects were asked to follow the gaze direction as precisely as they can. And if the gazer was
589 looking at a person, they were asked to respond if the target was present or absent. Each
590 participant finished sixteen practice trials to make sure they followed the instructions to follow
591 the gaze. During the practice, participants had unlimited time to familiarize themselves with
592 pictures of the target person. The videos in the practice session were different from the actual
593 experiment videos.

594
595 Participants then completed the main experimental sessions after practice trials. During a
596 session, observers completed all videos in random order. In total, each observer finished 2
597 sessions x 160 trials/session = 320 trials. Participants first finished a nine-point calibration and
598 validation. On each trial, the participants were instructed to fixate a cross and press the space
599 bar to start the trial. If the eye tracker detected an eye movement away from the fixation cross
600 of more than 1.5° visual angle when they pressed the space bar, the trial would not start, and

601 participants were required to recalibrate and revalidate. The cross was located exactly at the
602 location where the gazer's head would appear once the video started, so we can make sure the
603 participants were looking straight at the gazer and observing the gazing behavior. During the
604 video presentation (1.2s), participants were asked to follow the gaze direction as precisely as
605 they can. Once the video ended, the participants used a mouse to click if the target person was
606 present or not (Figure 1d).

607

608 AI model estimated gaze information

609 In order to quantify the gaze information in each video frame, we used a pre-trained deep
610 neural network (DNN) based model (Chong et al., 2020), which makes an objective estimate of
611 the gaze location for each video (Figure 3a-b). The model takes an entire image, a binary mask
612 that defines a bounding box around the gazer's head location, and a cropped image of the
613 gazer's head to produce a probability map of where the head's gaze is directed. We defined the
614 model gaze estimation as the pixel location corresponding with the highest probability on the
615 probability map. We repeated that for all the image frames from the video to obtain gaze
616 estimation over time. To estimate the head angular velocity, we first took the difference in
617 *angular displacement* for all continuous pairs of frames and smoothed the estimations by
618 convolving the differences with a kernel size of 5 frames. Similarly, we calculated the head
619 accelerations based on head velocity differences and smoothed them with a kernel size of 5.

620

621 Human estimated gaze information

622 Besides the AI model, we also recruited ten undergraduate research assistants to manually
623 annotate the gazer vectors for all the video frames where the target or distractor was digitally
624 erased. We used target/distractor-absent video frames for human annotations because we
625 want to use isolated gaze goal direction information based on the gazer head direction without
626 influences from peripheral information about potential gaze goals. We presented all the frames
627 in random order. Annotators used Matlab to click on each image to draw the estimated gazer
628 vector. We calculated the gazer vector angular error for each annotator and report the average
629 angular error as the final human estimated gazer vector for each frame.

630

631 Forward and Reverse Saccades Detection

632 We defined a *forward saccade* as an eye movement in which the direction vector had a positive
633 cosine similarity with the *gazer goal vector*. A *reverse saccade* was defined as a saccade vector
634 that happened after a forward saccade and had a negative cosine similarity with the *gazer goal*
635 *vector*. In addition, the reverse saccade endpoint was defined to be within a 2.5° visual angle
636 from the gazer to differentiate them from corrective saccades that overshoot the gazer goal. A
637 small subset of saccades was directed in the reverse direction because of the overshooting of the

638 gaze-following saccade. The endpoints of these reverse saccades had a mean of 6.7° distance to
639 the gazer's location. These saccades were considered different from reverse saccades to refixate
640 the gazer and were not included in the analysis.

641

642 Distribution of gaze information 200 ms before the gazer's head stops

643 In order to compute the general head velocity range before the gazer stops the gaze behaviors,
644 three annotators manually marked the time stamp when the gazer's head stops moving for
645 each movie independently. We then defined the gazer stops timing as the average time across
646 annotators for each movie. Finally, we calculated the mean gazer vector distance, angular
647 displacement, head velocity, and head acceleration during the range of 200ms right before the
648 gazer's head stops moving as the benchmarks (Figure 3c).

649

650 Statistical Analysis

651 We used within-subject ANOVA and t-tests for mean comparisons across different conditions.
652 We also used bootstrap techniques to estimate the statistical significance of variations of
653 saccade error (e.g., trials with reverse saccade vs. trials without reverse saccade) because of the
654 non-normality of the distributions. To apply the bootstrap test, we sampled 25 participants with
655 replacement and calculated the corresponding difference between conditions for each sampled
656 subject (a bootstrap sample), and repeated the process 10,000 times. The distribution of
657 resampled means or mean differences was used to assess statistical significance. All p values
658 were corrected using a false discovery rate (FDR) to reduce the probability of making a Type I
659 error. We used cluster-based permutation tests to compare the gazer's head velocity between
660 trials with reverse saccades and those without reverse saccades. We computed the mean
661 difference for each participant individually and permuted for 10,000 times. Based on
662 corrected p values, we acquired time intervals with significant differences. We used Python to
663 analyze all the data. For ANOVA tests we used package "pingouin" (Vallat, 2018). For the cluster-
664 based permutation test, we used the package "MNE" (Gramfort et al., 2013).

665

666 Support Vector Machine (SVM) Models

667 For training SVM models, we first computed the proportion of trials in which observers
668 executed reverse saccades for each movie. The median proportion of trials that included
669 reverse saccades for all movies was 20 %. We then did a median split of the movies into two
670 groups as a high possibility of the reverse saccade (>20%) vs. a low possibility of the reverse
671 saccade (<20%). Then we trained the SVM models with radial basis function kernel to classify
672 whether a movie had a high probability of triggering reverse saccade or not. We trained leave-
673 one-out SVM based on four gazer vector features: 1. Gazer vector distance 2. Gaze angular
674 displacement 3. Head angular velocity 4. Head angular acceleration. For training each SVM
675 model, we chose one of the four gaze features during a specific time range from the video

676 onset as the predictor. We used the package "sklearn" for the training process (Pedregosa et
677 al., 2011).

678

679 *Experiment 1A (random freeze)*

680 Subjects

681 Five undergraduate students (ages from 18-20, 2 male, 3 female) from the University of
682 California Santa Barbara were recruited as subjects for credits in this experiment. All have
683 normal to corrected-to-normal vision. All participants signed consent forms to participate in the
684 study.

685

686 Experimental Setup and Stimuli

687 We had the same experiment stimuli and setup as experiment 1, except that we detected
688 saccades during the movie presentation. When we detected the first gaze-following saccade,
689 there was a 50% chance the gazer would be frozen (without movement) for the rest of the movie
690 to prevent any motion afterward that could potentially attract reverse saccade.

691

692 Procedure

693 The procedure was the same as experiment 1. Participants were told to follow the gaze as
694 precisely as they could during the movie presentation. In total, each observer finished 2 sessions
695 x 160 trials/session = 320 trials. And participants were not aware of the random freezing of the
696 video.

697

698 *Experiment 2*

699 Subjects

700 Twenty-five undergraduate students (ages from 18-20, 10 male, 15 female) from the University
701 of California Santa Barbara were recruited as subjects for credits in this experiment. All had
702 normal to corrected-to-normal vision. All participants signed consent forms to participate in the
703 study.

704

705 Experimental Setup and Stimuli

706 We had the same experiment stimuli and setup as experiment 1, except that we detected reverse
707 saccade during the movie presentation. When we detected a reverse saccade back to the gazer
708 after the first gaze-following saccade, there was a 50% chance the gazer would be completely
709 erased for the rest of the movie, to prevent any foveal processing of the gazer.

710

711

712 Procedure

713 The procedure was the same behavioral task as experiment 1. Participants were told to follow
714 the gaze as precisely as they could during the movie presentation. And participants were unaware
715 of the random erasure of the gazer.

716

717 *Experiment 3 (free-viewing search)*

718 Subjects

719 Five undergraduate students (ages from 18-21, 3 male, 2 female) from the University of
720 California Santa Barbara were recruited as subjects for credits in this experiment. All have
721 normal to corrected-to-normal vision. All participants signed consent forms to participate in the
722 study.

723

724 Experimental setup and stimuli

725 We used the same experiment stimuli and setup as in experiment 1.

726

727 Procedure

728 The procedure was the same as experiment 1, except that we did not instruct participants to
729 follow the gaze explicitly. Instead, we asked them to just free-viewing the video and respond
730 whether the target person was present or absent. In total, each observer finished 2 sessions x
731 160 trials/session = 320 trials.

732

733 **Competing Interests**

734 The authors declare that they have no competing interests.

735

736 **Data Availability**

737 Video Stimuli and Data that support the findings of this study have been deposited at osf

738 <https://osf.io/g9bzt/>.

739

740 **Code Availability**

741 Code to replicate analysis is available at osf: <https://osf.io/g9bzt/>.

742

743 **Contribution**

744 NXH and MPE conceived and designed the studies and determined the data analysis. NXH
745 implemented the experiment, collected and analyzed the data. NXH and MPE both wrote the
746 manuscript. Both authors approved the final version of the manuscript for submission.

747 **Acknowledgement**

748 The research was sponsored by the U.S. Army Research Office and was accomplished under
749 Contract Number W911NF-19-D-0001 for the Institute for Collaborative Biotechnologies. MPE
750 was supported by a Guggenheim Foundation Fellowship. The views and conclusions contained
751 in this document are those of the authors and should not be interpreted as representing the
752 official policies, either expressed or implied, of the U.S. Government. The U.S. Government is
753 authorized to reproduce and distribute reprints for Government purposes notwithstanding any
754 copyright notation herein.

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777 **References**

- 778 Araujo, C., Kowler, E., & Pavel, M. (2001). Eye movements during visual search: The costs of
779 choosing the optimal path. *Vision Research*, 41(25–26), 3613–3625.
- 780 Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye
781 movements. *Mathematical Biosciences*, 24(3), 191–204. [https://doi.org/10.1016/0025-](https://doi.org/10.1016/0025-5564(75)90075-9)
782 5564(75)90075-9
- 783 Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory Representations in Natural Tasks.
784 *Journal of Cognitive Neuroscience*, 7(1), 66–80.
785 <https://doi.org/10.1162/jocn.1995.7.1.66>
- 786 Balsdon, T., & Clifford, C. (2018). Task Dependent Effects of Head Orientation on Perceived
787 Gaze Direction. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.02491>
- 788 Baron-Cohen, S. (2001). *Mindblindness: An essay on autism and theory of mind*. MIT Press.
- 789 Basu, D., & Murthy, A. (2020). Parallel programming of saccades in the macaque frontal eye
790 field: Are sequential motor plans coactivated? *Journal of Neurophysiology*, 123(1), 107–
791 119. <https://doi.org/10.1152/jn.00545.2018>
- 792 Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step
793 stimuli. *Vision Research*, 19(9), 967–983. [https://doi.org/10.1016/0042-6989\(79\)90222-](https://doi.org/10.1016/0042-6989(79)90222-0)
794 0
- 795 Bekerman, I., Gottlieb, P., & Vaiman, M. (2014). Variations in eyeball diameters of the healthy
796 adults. *Journal of Ophthalmology*, 2014, 503645. <https://doi.org/10.1155/2014/503645>

- 797 Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to
798 language. *Developmental Science*, 8(6), 535–543. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-7687.2005.00445.x)
799 [7687.2005.00445.x](https://doi.org/10.1111/j.1467-7687.2005.00445.x)
- 800 Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition,
801 Joint Attention, and Communicative Competence from 9 to 15 Months of Age.
802 *Monographs of the Society for Research in Child Development*, 63(4), i–174.
803 <https://doi.org/10.2307/1166214>
- 804 Caspi, A., Beutter, B. R., & Eckstein, M. P. (2004). The time course of visual information accrual
805 guiding eye movement decisions. *Proceedings of the National Academy of Sciences of*
806 *the United States of America*, 101(35), 13086–13090.
807 <https://doi.org/10.1073/pnas.0305329101>
- 808 Chawarska, K., Klin, A., & Volkmar, F. (2003). Automatic attention cueing through eye
809 movement in 2-year-old children with autism. *Child Development*, 74(4), 1108–1122.
810 <https://doi.org/10.1111/1467-8624.00595>
- 811 Chong, E., Wang, Y., Ruiz, N., & Rehg, J. M. (2020). Detecting Attended Visual Targets in Video.
812 *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5395–
813 5405. <https://doi.org/10.1109/CVPR42600.2020.00544>
- 814 Cline, M. G. (1967). The perception of where a person is looking. *The American Journal of*
815 *Psychology*, 80(1), 41–50.
- 816 Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-cohen, S. (1999). Gaze
817 perception triggers reflexive visuospatial orienting. *Visual Cognition*, 509–540.

- 818 Eckstein, M. P., Schoonveld, W., Zhang, S., Mack, S. C., & Akbas, E. (2015). Optimal and human
819 eye movements to clustered low value cues to increase decision rewards during search.
820 *Vision Research*, *113*, 137–154.
- 821 Egeth, H. E., & Yantis, S. (1997). VISUAL ATTENTION: Control, Representation, and Time Course.
822 *Annual Review of Psychology*, *48*(1), 269–297.
823 <https://doi.org/10.1146/annurev.psych.48.1.269>
- 824 Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze.
825 *Neuroscience & Biobehavioral Reviews*, *24*(6), 581–604. [https://doi.org/10.1016/S0149-](https://doi.org/10.1016/S0149-7634(00)00025-7)
826 [7634\(00\)00025-7](https://doi.org/10.1016/S0149-7634(00)00025-7)
- 827 Fooker, J., & Spering, M. (2020). Eye movements as a readout of sensorimotor decision
828 processes. *Journal of Neurophysiology*, *123*(4), 1439–1447.
829 <https://doi.org/10.1152/jn.00622.2019>
- 830 Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by
831 nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490–495.
832 <https://doi.org/10.3758/BF03208827>
- 833 Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and
834 arrow cues. *Journal of Experimental Psychology. Human Perception and Performance*,
835 *30*(2), 319–329. <https://doi.org/10.1037/0096-1523.30.2.319>
- 836 Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M.,
837 Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with
838 MNE-Python. *Frontiers in Neuroscience*, *7*.
839 <https://www.frontiersin.org/articles/10.3389/fnins.2013.00267>

- 840 Gregory, S. (2021). Investigating facilitatory versus inhibitory effects of dynamic social and non-
841 social cues on attention in a realistic space. *Psychological Research*.
842 <https://doi.org/10.1007/s00426-021-01574-7>
- 843 Han, N. X., & Eckstein, M. P. (2022). Gaze-cued shifts of attention and microsaccades are
844 sustained for whole bodies but are transient for body parts. *Psychonomic Bulletin &*
845 *Review*.
- 846 Han, N. X., Wang, W. Y., & Eckstein, M. P. (2021). Gaze Perception in Humans and CNN-Based
847 Model. *ArXiv:2104.08447 [Cs]*. <http://arxiv.org/abs/2104.08447>
- 848 Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive*
849 *Sciences*, 9(4), 188–194. <https://doi.org/16/j.tics.2005.02.009>
- 850 Hayhoe, M. M. (2017). Vision and Action. *Annual Review of Vision Science*, 3(1), 389–413.
851 <https://doi.org/10.1146/annurev-vision-102016-061437>
- 852 Hooge, I. T., & Erkelens, C. J. (1999). Peripheral vision and oculomotor control during visual
853 search. *Vision Research*, 39(8), 1567–1575.
- 854 Hoppe, D., & Rothkopf, C. A. (2019). Multi-step planning of eye movements in visual search.
855 *Scientific Reports*, 9(1), 144. <https://doi.org/10.1038/s41598-018-37536-0>
- 856 Inhoff, A. W., Kim, A., & Radach, R. (2019). Regressions during Reading. *Vision*, 3(3), 35.
857 <https://doi.org/10.3390/vision3030035>
- 858 Intoy, J., & Rucci, M. (2020). Finely tuned eye movements enhance visual acuity. *Nature*
859 *Communications*, 11(1), 795. <https://doi.org/10.1038/s41467-020-14616-2>
- 860 Jonides, J., & Jonides, J. (1981). *Voluntary versus automatic control over the mind's eye's*
861 *movement*.

- 862 [http://wexler.free.fr/library/files/jonides%20\(1981\)%20voluntary%20versus%20automa](http://wexler.free.fr/library/files/jonides%20(1981)%20voluntary%20versus%20automa)
863 [tic%20control%20over%20the%20mind's%20eye's%20movements.pdf](http://wexler.free.fr/library/files/jonides%20(1981)%20voluntary%20versus%20automa)
- 864 Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2005). Domestic goats, *Capra hircus*, follow
865 gaze direction and use social cues in an object choice task. *Animal Behaviour*, *69*(1), 11–
866 18. <https://doi.org/10.1016/j.anbehav.2004.05.008>
- 867 Kingstone, A., Smilek, D., Ristic, J., Kelland Friesen, C., & Eastwood, J. D. (2003). Attention,
868 researchers! It is time to take a look at the real world. *Current Directions in*
869 *Psychological Science*, *12*(5), 176–180.
- 870 Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*(1),
871 78–100. <https://doi.org/10.1037/0033-2909.100.1.78>
- 872 Kowler, E. (1989). Cognitive expectations, not habits, control anticipatory smooth oculomotor
873 pursuit. *Vision Research*, *29*(9), 1049–1057. <https://doi.org/10.1016/0042->
874 [6989\(89\)90052-7](https://doi.org/10.1016/0042-6989(89)90052-7)
- 875 Kowler, E. (2011). Eye movements: The past 25years. *Vision Research*, *51*(13), 1457–1483.
876 <https://doi.org/10.1016/j.visres.2010.12.014>
- 877 Kowler, E., Rubinstein, J. F., Santos, E. M., & Wang, J. (2019). Predictive Smooth Pursuit Eye
878 Movements. *Annual Review of Vision Science*, *5*, 223–246.
879 <https://doi.org/10.1146/annurev-vision-091718-014901>
- 880 Langton, S. R. H. (2000). The mutual influence of gaze and head orientation in the analysis of
881 social attention direction. *The Quarterly Journal of Experimental Psychology Section A*,
882 *53*(3), 825–845. <https://doi.org/10.1080/713755908>

- 883 Langton, S. R. H., & Bruce, V. (1999). Reflexive Visual Orienting in Response to the Social
884 Attention of Others. *Visual Cognition*, 6(5), 541–567.
885 <https://doi.org/10.1080/135062899394939>
- 886 Langton, S. R. H., Honeyman, H., & Tessler, E. (2004). The influence of head contour and nose
887 angle on the perception of eye-gaze direction. *Perception & Psychophysics*, 66(5), 752–
888 771. <https://doi.org/10.3758/BF03194970>
- 889 Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of
890 social attention. *Trends in Cognitive Sciences*, 4(2), 50–59.
891 [https://doi.org/10.1016/S1364-6613\(99\)01436-9](https://doi.org/10.1016/S1364-6613(99)01436-9)
- 892 Lanzilotto, M., Gerbella, M., Perciavalle, V., & Lucchetti, C. (2017). Neuronal Encoding of Self
893 and Others' Head Rotation in the Macaque Dorsal Prefrontal Cortex. *Scientific Reports*,
894 7(1), 8571. <https://doi.org/10.1038/s41598-017-08936-5>
- 895 Laubrock, J., Cajar, A., & Engbert, R. (2013). Control of fixation duration during scene viewing by
896 interaction of foveal and peripheral processing. *Journal of Vision*, 13(12), 11.
897 <https://doi.org/10.1167/13.12.11>
- 898 Lee, J.-H., Hwang Shin, S.-J., & Istook, C. L. (2006). Analysis of Human Head Shapes in the United
899 States. *International Journal of Human Ecology*, 7(1), 77–83.
- 900 Leekam, S. R., Hunnisett, E., & Moore, C. (1998). Targets and Cues: Gaze-following in Children
901 with Autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 39(7),
902 951–962. <https://doi.org/10.1111/1469-7610.00398>

- 903 Legge, G. E., Hooven, T. A., Klitz, T. S., Stephen Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips
904 2002: New insights from an ideal-observer model of reading. *Vision Research*, *42*(18),
905 2219–2234.
- 906 Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: An ideal-observer model of reading.
907 *Psychological Review*, *104*(3), 524–553. <https://doi.org/10.1037/0033-295X.104.3.524>
- 908 Ludwig, C. J. H., Davies, J. R., & Eckstein, M. P. (2014). Foveal analysis and peripheral selection
909 during active visual sampling. *Proceedings of the National Academy of Sciences*,
910 201313553. <https://doi.org/10.1073/pnas.1313553111>
- 911 Ludwig, C. J. H., Gilchrist, I. D., McSorley, E., & Baddeley, R. J. (2005). The Temporal Impulse
912 Response Underlying Saccadic Decisions. *Journal of Neuroscience*, *25*(43), 9907–9912.
913 <https://doi.org/10.1523/JNEUROSCI.2197-05.2005>
- 914 Mansfield, E., Farroni, T., & Johnson, M. (2003). Does gaze perception facilitate overt orienting?
915 *Visual Cognition*, *10*(1), 7–14. <https://doi.org/10.1080/713756671>
- 916 Meghanathan, R. N., Nikolaev, A. R., & van Leeuwen, C. (2019). Refixation patterns reveal
917 memory-encoding strategies in free viewing. *Attention, Perception & Psychophysics*,
918 *81*(7), 2499–2516. <https://doi.org/10.3758/s13414-019-01735-2>
- 919 Mennie, N., Hayhoe, M., & Sullivan, B. (2007). Look-ahead fixations: Anticipatory eye
920 movements in natural tasks. *Experimental Brain Research. Experimentelle*
921 *Hirnforschung. Expérimentation Cérébrale*, *179*(3), 427–442.
922 <https://doi.org/10.1007/s00221-006-0804-0>
- 923 Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K.
924 (2000). Responding to Joint Attention Across the 6- Through 24-Month Age Period and

- 925 Early Language Acquisition. *Journal of Applied Developmental Psychology*, 21(3), 283–
926 298. [https://doi.org/10.1016/S0193-3973\(99\)00040-4](https://doi.org/10.1016/S0193-3973(99)00040-4)
- 927 Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language
928 development in 6-month-olds. *Infant Behavior and Development*, 21(2), 373–377.
929 [https://doi.org/10.1016/S0163-6383\(98\)90014-5](https://doi.org/10.1016/S0163-6383(98)90014-5)
- 930 Mulckhuyse, M., & Theeuwes, J. (2010). Unconscious attentional orienting to exogenous cues: A
931 review of the literature. *Acta Psychologica*, 134(3), 299–309.
932 <https://doi.org/10.1016/j.actpsy.2010.03.002>
- 933 Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*,
934 434(7031), 387–391. <https://doi.org/10.1038/nature03390>
- 935 Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is
936 it social? *Development and Psychopathology*, 20(1), 79–97.
937 <https://doi.org/10.1017/S0954579408000047>
- 938 Or, C. C.-F., Peterson, M. F., & Eckstein, M. P. (2015). Initial eye movements during face
939 identification are optimal and similar across cultures. *Journal of Vision*, 15(13), 12–12.
940 <https://doi.org/10.1167/15.13.12>
- 941 Oram, M. W., & Perrett, D. I. (1994). Responses of Anterior Superior Temporal Polysensory
942 (STPa) Neurons to “Biological Motion” Stimuli. *Journal of Cognitive Neuroscience*, 6(2),
943 99–116. <https://doi.org/10.1162/jocn.1994.6.2.99>
- 944 Otsuka, Y. (2014). *Dual-route model of the effect of head orientation on perceived gaze*
945 *direction*. *Journal of Experimental Psychology: Human Perception and Performance*; US:
946 American Psychological Association. <https://doi.org/10.1037/a0036151>

- 947 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M.,
948 Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D.,
949 Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in
950 Python. *The Journal of Machine Learning Research*, 12(null), 2825–2830.
- 951 Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex
952 tasks. *Vision Research*, 41(25–26), 3587–3596. [https://doi.org/10.1016/s0042-](https://doi.org/10.1016/s0042-6989(01)00245-0)
953 [6989\(01\)00245-0](https://doi.org/10.1016/s0042-6989(01)00245-0)
- 954 Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face
955 recognition tasks. *Proceedings of the National Academy of Sciences of the United States*
956 *of America*, 109(48), E3314–3323. <https://doi.org/10.1073/pnas.1214269109>
- 957 Pi, Z., Xu, K., Liu, C., & Yang, J. (2020). Instructor presence in video lectures: Eye gaze matters,
958 but not body orientation. *Computers & Education*, 144, 103713.
959 <https://doi.org/10.1016/j.compedu.2019.103713>
- 960 Pierrot-Deseilligny, C., Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral
961 cortex. *Current Opinion in Neurology*, 17(1), 17–25. [https://doi.org/10.1097/00019052-](https://doi.org/10.1097/00019052-200402000-00005)
962 [200402000-00005](https://doi.org/10.1097/00019052-200402000-00005)
- 963 Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research.
964 *Psychological Bulletin*, 124(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- 965 Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements
966 reduce local uncertainty. *Journal of Vision*, 7(3), 6. <https://doi.org/10.1167/7.3.6>
- 967 Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at
968 it. *Psychonomic Bulletin & Review*, 9(3), 507–513. <https://doi.org/10.3758/bf03196306>

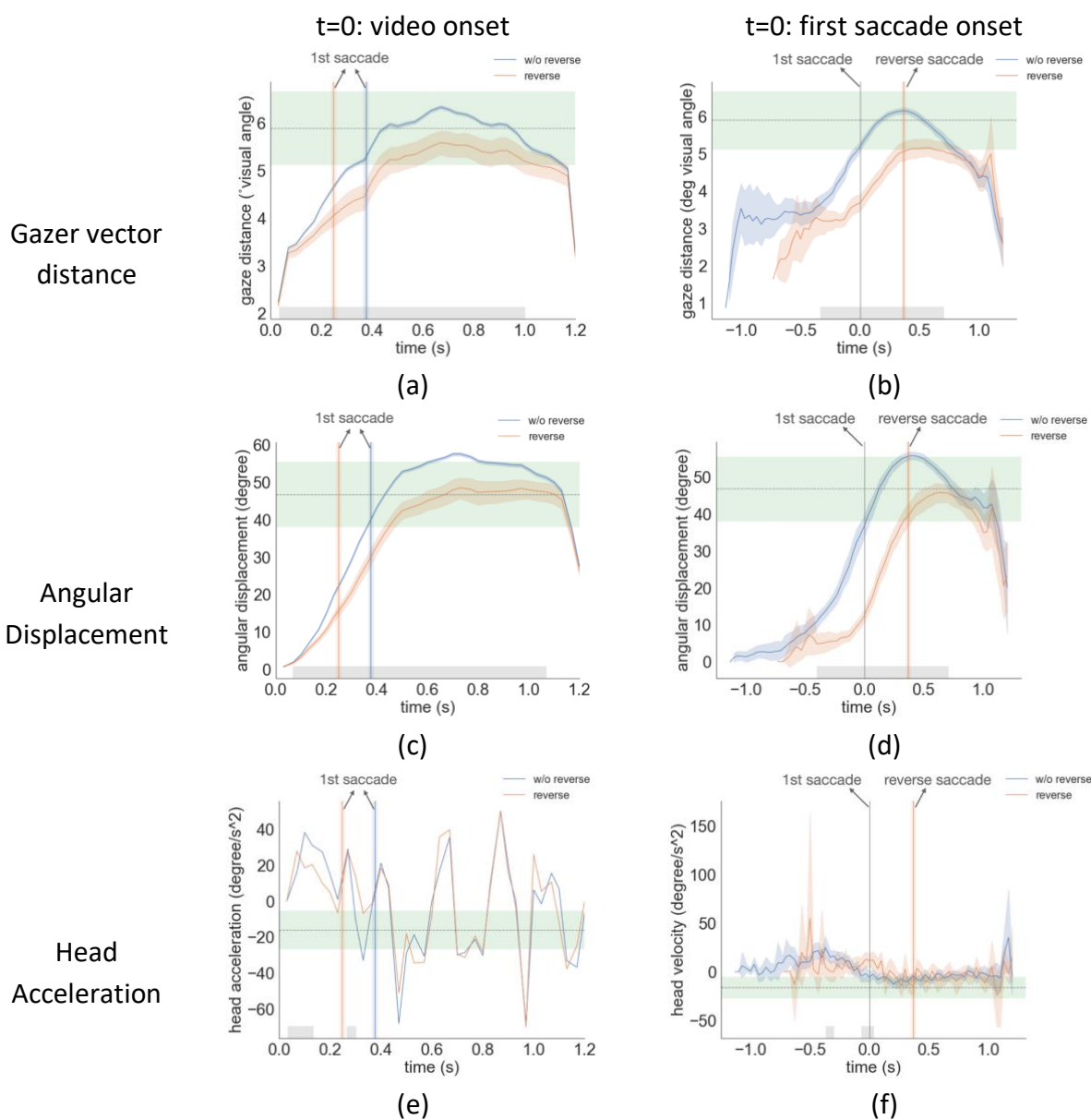
- 969 Ristic, J., Mottron, L., Friesen, C. K., Iarocci, G., Burack, J. A., & Kingstone, A. (2005). Eyes are
970 special but not for everyone: The case of autism. *Brain Research. Cognitive Brain*
971 *Research*, 24(3), 715–718. <https://doi.org/10.1016/j.cogbrainres.2005.02.007>
- 972 Rucci, M., Iovin, R., Poletti, M., & Santini, F. (2007). Miniature eye movements enhance fine
973 spatial detail. *Nature*, 447(7146), 852–855. <https://doi.org/10.1038/nature05866>
- 974 Senju, A., & Csibra, G. (2008). Gaze Following in Human Infants Depends on Communicative
975 Signals. *Current Biology*, 18(9), 668–671. <https://doi.org/10.1016/j.cub.2008.03.059>
- 976 Shepherd, S. V. (2010). Following Gaze: Gaze-Following Behavior as a Window into Social
977 Cognition. *Frontiers in Integrative Neuroscience*, 4.
978 <https://doi.org/10.3389/fnint.2010.00005>
- 979 Stewart, E. E. M., Valsecchi, M., & Schütz, A. C. (2020). A review of interactions between
980 peripheral and foveal vision. *Journal of Vision*, 20(12), 2.
981 <https://doi.org/10.1167/jov.20.12.2>
- 982 Vallat, R. (2018). Pingouin: Statistics in Python. *Journal of Open Source Software*, 3(31), 1026.
983 <https://doi.org/10.21105/joss.01026>
- 984 Wallis, L. J., Range, F., Müller, C. A., Serisier, S., Huber, L., & Virányi, Z. (2015). Training for eye
985 contact modulates gaze following in dogs. *Animal Behaviour*, 106, 27–35.
986 <https://doi.org/10.1016/j.anbehav.2015.04.020>
- 987 Wang, L., Yang, X., Shi, J., & Jiang, Y. (2014). The feet have it: Local biological motion cues
988 trigger reflexive attentional orienting in the brain. *NeuroImage*, 84, 217–224.
989 <https://doi.org/10.1016/j.neuroimage.2013.08.041>

- 990 Wolf, C., Belopolsky, A. V., & Lappe, M. (2022). Current foveal inspection and previous
991 peripheral preview influence subsequent eye movement decisions. *iScience*, 25(9),
992 104922. <https://doi.org/10.1016/j.isci.2022.104922>
- 993 Wolf, C., & Schütz, A. C. (2015). Trans-saccadic integration of peripheral and foveal feature
994 information is close to optimal. *Journal of Vision*, 15(16), 1.
995 <https://doi.org/10.1167/15.16.1>
- 996 Woodward, A. L. (2003). Infants' developing understanding of the link between looker and
997 object. *Developmental Science*, 6(3), 297–311. [https://doi.org/10.1111/1467-](https://doi.org/10.1111/1467-7687.00286)
998 7687.00286

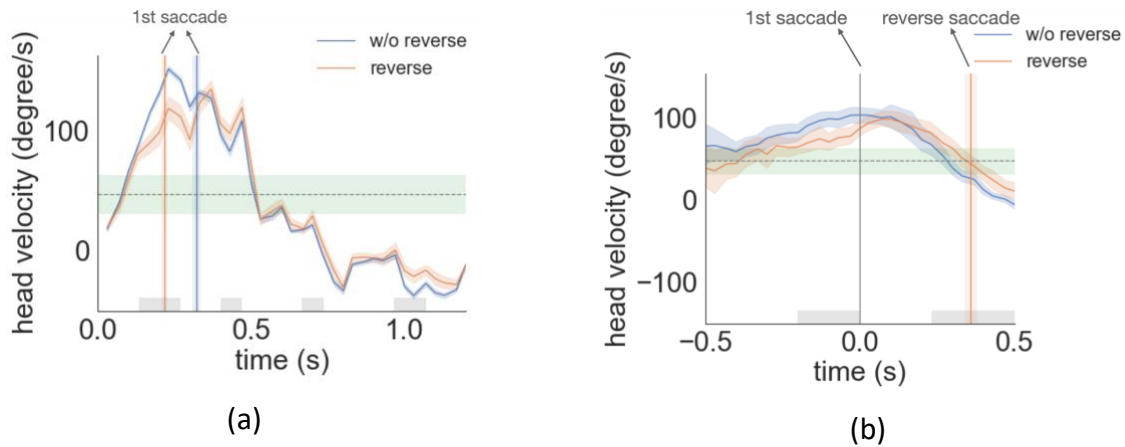
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020

1021 **Supplementary Material**

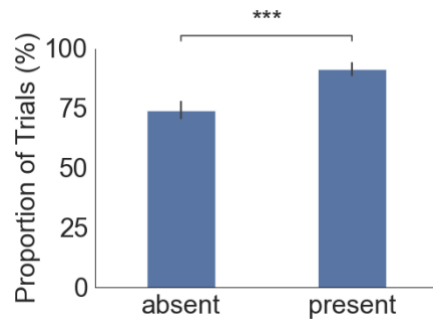
1022



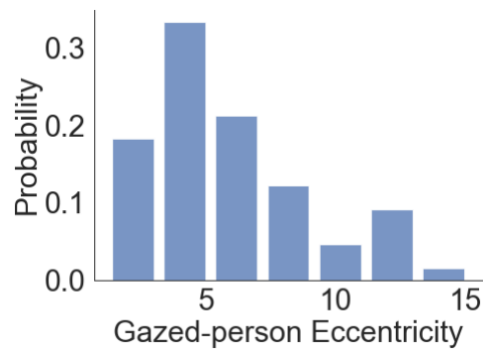
S1. Gazer vector distance, angular displacement, and head acceleration over time, with t=0 aligned with video onset (left column) and t=0 aligned with first saccade onset (right column). The shaded area was the 95% bootstrapped confidence interval. The gray area was significant under the cluster-based permutation test. The green area was the 95% confidence interval of each gaze variable right before the gazer's head stops moving across all movies.



S2. Gazer's head velocity from the beginning of the movie (a) and aligned with the first saccade initiation time (b) for experiment 2.



S3. The proportion of trials that have the first saccade following the gaze direction for the *absent* and the *present* condition.



S4. The histogram of gazed person eccentricities relative to the gazer's position.