PERCEPTION OF GAZE AND POINTING IN DYNAMIC ENVIRONMENT

Ritu Moni Borah

A dissertation submitted to the faculty of
National Brain Research Centre
in partial fulfillment of the requirements for
Masters in Neuroscience

National Brain Research Centre
Manesar, Gurugram, Haryana
May 2018

CERTIFICATE

This is to certify that the dissertation entitled “Perception of gaze and pointing in dynamic environment” was carried out by Ritu Moni Borah at National Brain Research Centre, Manesar, Haryana, as partial fulfillment for the MSc degree.

The work presented herein is original and has not been submitted previously for the award of any degree or diploma at National Brain Research Centre (Deemed University) or any other University.

Supervisor   Director

Place: NBRC, Manesar

Date: May, 2018

DECLARATION BY THE CANDIDATE

I Ritu Moni Borah hereby declare that the work presented in this dissertation is carried out by me, under the guidance of Dr. Arpan Banerjee, National Brain Research Centre, Manesar, Haryana.

I also declare that no part of this dissertation has been previously submitted for the award of any degree or diploma at National Brain Research Centre or any other University.

Supervisor   Ritu Moni Borah

Place:

Date:

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my advisor Dr. Arpan Banerjee for continuous support and dedication throughout the dissertation work. Without his guidance this work would not have been accomplished.

I would like to thank Nilambari Hajare, Vinodh G Kumar, Neeraj Kumar, Moumita Das, Amit Ranjan, Dr Shyamchand Soibam, Bikas Sahoo who taught me various tools and methods. And also thanks for the introspective discussions which helped me a lot to interpret my own work critically.

I would also like to thank to Dr.Dipanjan Ray, Shrey Dutta, Dr.Amit Naskar, Siddharta Talwar, Priyanka Ghosh, Anagh Pathak for being supportive and sharing valuable knowledge.

My heartfelt thanks to Tripti Joshi for her volunteering to help out with experimental design and being supportive and understanding. And thanks for being patient with the video shoots. Without her help, the work would not have been completed.

Also special thanks to my parents for being understanding and supportive and thanks to Nilambari and Tripti again for encouraging chats and refreshing ideas.

ABSTRACT

Author: Ritu Moni Borah

Title: Perception of gaze and point in dynamic environment

Institution: National Brain Research Centre

Dissertation Advisor: Dr Arpan Banerjee

Degree: Masters in Neuroscience

Year: 2018

Gaze and pointing are very common and important social cues used in our everyday routines, for example to draw attention of a person toward an intended object. Earlier studies on gaze and pointing perception have not explored in detail about the integration of information in dynamical context. Our study focuses on understanding the correlates of eye gaze and pointing with gaze behavior in a dynamic environment when both stimuli are presented together. We are particularly interested to study how the integration takes place when the gaze and pointing information are provided with uncertainity in time. Conflicting stimuli were also presented to evaluate which cue source is more reliable to predict the interlocutor's intention or reference.

CONTENTS

1. INTRODUCTION

Pointing and gazing are very commonly used social gestures to relay our communicative intentions to others,for example, to draw attention towards an intended object or a person.

Grasping and reaching conveys intention of an agent unambiguously, however, to an observer pointing with an index finger may be interpreted as a hand–object interaction (e.g. referring to the object to just draw attention) or an indication to interact with the object (e.g. to grasp or pass the object to the agent) [20]. Similarly, eye movements can be used to direct someone's attention [24] or understand others' intention to act upon objects [22,1]. Thus, gaze and pointing cues are important factors for conveying motor or communicative intentions. Moreover, gaze and pointing may occur together in some context where the observer will have to perceive multiple visual information to comprehend the agent's goals or intentions [4,10]. This suggests the need to study both these cues together.

The current study aims to advance the understanding that how gaze and pointing are perceived in dynamic environment. When gaze and pointing start time is uncertain, how does that affect our perception? So, what are the consequences when both cues start together i.e. zero lag of movement's start, and when there is a lag? Finally how is the information integration taking place in these scenarios?

To achieve our aim, we designed a paradigm with stimuli videos showing cue (gaze/pointing or both) and object interaction. Cue and object interactions were shown, because gaze shift or pointing usually occur with respect to some reference to an object or person and not alone. So, stimuli videos contained an actor performing either gaze shift or pointing or both (conflicting or congruent) toward an object. Each video had different time of actor's gaze onset and also there was lag with start of hand movement.

Earlier studies have focused on perception of gaze and pointing in static environment (e.g., [4, 17, 21, 24]), and the effect of dynamic environment has not been explored in much detail. In a relevant study (considering gaze and motor cue in dynamic environment), *Ambrosini et al. 2015* investigated the integration of information by providing fixated gaze of an actor at an object, with the hand pre-shape in the next and then the arm movement was provided. Their result suggested that gaze information is a reliable source for predicting others' intentions only when no other information about the actor's behavioral intention was provided. Moreover, they found that hand pre-shape information was considered for action predictions by the participants even when that source was made unreliable (i.e. conflicting conditions). This implies that motor cues are more informative in competing environment.

Assuming all kind of motor cues will effectively be dominant in competing environment as demonstrated by series of experiment by *Ambrosini et al. 2011,2015*, we also expect that pointing would also generate the similar behavioral response.

Predictive gaze shifts are important in guiding our own actions and also integral to observing another person's goal directed action [7,10]. This contribute to direct matching hypothesis which suggests that eye movements produced while performing a task is same when observing the task, benefited by the observer's own motor knowledge [7] .This prediction is required for successful interactions with the environment, without which perception of the surrounding world would be fragmented and lag with the real world events will be produced, irrespective of goal being collaborative or competitive [10].

Thus, observer's gaze will be at the referred objects before the actor's action completes. We also predict that integration of information will be faster in congruent conditions compared to incongruent and the gaze of the observer will take lead to that of the actor's movement completion irrespective of congruent or incongruent condition. An increase should be seen with lags (within conditions) , in observer's gaze reaching time to the referred object.

So questions addressed in this dissertation are:

A. How gaze shifts are spatially distributed in provided conditions?
B. How across conditions, the time of gaze arrival at the referred object differs?

2. EYE MOVEMENTS

Throughout the animal kingdom, consistent pattern of ‘saccade and fixate’ strategy of eye movements are reported [15,16]. Saccades are the rapid movements that redirect the eye to a new part of the surroundings, and fixations are intervals between saccades in which gaze state is stationary [19, 13, 15]. These movements are important to maintain the image of the target onto the fovea which is a very small region at the retina densely packed with cones (photoreceptor).

To maintain the focus on target of interest onto fovea, relocations through saccades have to be very fast. The reaching speeds of larger saccades can be upto 700 °/s, thus we are effectively blind during these periods [15]. However, gaze must be still between saccades i.e. during fixations since ~20 ms are required for cones to respond effectively to a step change in the light reaching it [8, 15].

Apart from saccade and fixation, two other eye movements are observed: smooth pursuit and vergence (for details refer to [8, 13, 15]).

These eye movements are not just some random movements, but depict the observers’ goal directed search or intention or a salient feature of a stimuli [19]. Thus, studies of gaze location, especially saccades and fixations, are widely studied to understand the underlying cognitive processes related to the task specific interaction. This trend was set by Alfred Yarbus, who was the first to demonstrate that eye movements shows attentional arrest of a stimuli and also defines strategy that an observer use for performing a task [19].

Thus, gaze being goal-directed will tell us about the decisions made or strategies followed by an observer for the a presented stimuli. Hence, we aim to investigate fixational gaze for our study.

2.1 Eye Tracking

Eye movements can be measured by the means of eye tracker which now-a-days work on video based measurement. The eye is illuminated with infrared source and the resulting corneal reflection (glint or 1st Purkinje reflection) or pupil reflection (4th Purkinje reflection) is recorded. The corneal reflection is the brightest but no the only. Some trackers record make use of both the corneal and pupil reflection to test reliability of their data [13].

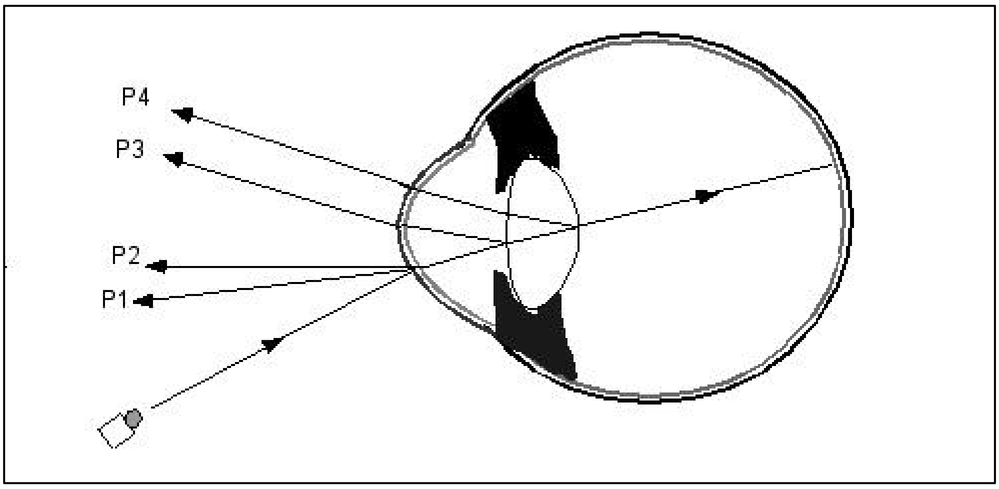


Figure 1: the four Purkinje images or reflections (source: http://www.opt.indiana.edu/v665/CD/CD_Version/CH4/CH4.HTM

2.2 Eye Movements in Action Observation

Evidence of action observation being predictive comes from the eye movements recording [3]. *Flanagan & Johansson 2003* demonstrated that people make same kind of eye movements when they perform a task themselves (block stacking in their experiment) and when they are viewing that task. They proposed that task specific proactive eye movements may be linked to neural processes for planning and control of that manual action. This was consistent with mirror neurons system which works on visuo-motor matching process to elicit predictive response and consequently, suggesting action understanding relies on motor system [3]. These findings were further supported, for example, by transcranial magnetic stimulation (TMS) studies done by *Elsner et al. 2013, Constantini et al. 2013,* and *Constantini et al. 2014*. TMS pulses targeting ventral premotor cortex or superior temporal sulcus or frontal eye field were subjected while observers were performing the task and as expected this resulted in delayed goal directed gaze shift compared to the control conditions in their respective studies. Moreover, developmental studies have suggested that prediction of action is linked with action performing ability, statistical regularity or repetition of the action, and goal’s saliency [3].

3. GAZE AND POINTING

Perception of gaze has been shown to recruit anterior and posterior fusiform gyrus, bilateral superior temporal sulcus (STS), left anterior STS, intra-parietal sulcus/superior parietal lobule, postcentral gyrus, and right inferior frontal gyrus [11].

Materna et al. 2008 provided evidence that posterior STS was recruited for directional perception of socially relevant cues which included gaze and pointing too. This is required for re-direction of one's attentional focus toward the cue. *Sato et al. 2009* showed commonalities in attention triggered shift due to directional hand, gaze, and symbols. Right inferior parietal lobule, inferior frontal gyrus, STS, occipital cortices responded to directional cues than non-directional. Moreover, in response to directional gaze cue, amygdala was also activated which they reported was in line with previous studies of lesion amygdala depicting impaired attentional shift to gaze.

Pierno et al. 2009 showed that visual perception of hand (in still image) triggered BOLD response in the lateral occipitotemporal cortex and elicited response was seen with grasping or pointing actions in the same region. They also showed that observing someone's hand grasping an object activated the bilateral somatosensory cortex, even when the hands shown were relaxed and immobile. This suggest that the mere sight of pointing induces an intent on observer to act upon the referred object. Moreover, their results revealed that the perception of resting hand, non-interacting with the placed object, also triggers the same response as a grasping and pointing hand. These led to activation of parietal region and dorsal and ventral premotor cortex. *Conty et al. 2012* showed that right parietal and supplementary motor cortices were recruited for integration of someone's pointing with gaze direction.

4. METHODS

4.1 Participants

Seventeen right handed volunteers (female = 7, male = 10) within the age group of 21-27 years (mean=24.31) participated in the study. One of the participant's data was excluded due to deviant fixations. The participants were untrained with respect to the purpose of the study.

4.2 Ethics Statement

Written consent was obtained from each participant prior to the experiments. The experiments were approved by the ethics committee of the institute.

4.3 Stimuli and Apparatus

The paradigm presented by Peeters et al 2017 [10] was amended to fit our question of interest. Stimuli contained of trial videos, wherein an instructor was performing an action of directing gaze or point towards an object which were either left or right directed. Four conditions were provided which includes "Congruent Condition" where the instructor was looking and pointing at an object in the same direction, "Incongruent Condition" where the instructor was gazing and pointing at different object. There were also "OnlyGaze" and "OnlyPoint" trials as control. Within the Congruent and Incongruent condition, lagged and synchronous videos were presented. The details of stimuli timing are explained in Table 1. There were a total of sixteen different videos presented with combination of six different elongated objects (each repeated 25 times) with 50% distribution of referred object in either left or right direction, exception being in incongruent trials where reference to both left and right objects were there. Trial videos (1.579±0.376 s) were presented in randomization with fixation cross pictures (duration variability of 1.2983±0.417 s) as inter-trial interval. The videos start with instructor sitting neutrally, gazing in front and then with variability of 172.22±16.87 ms starts the action. Pointing instructions presented by the instructor were with the right hand which is the dominant hand used in most cases.

		G	G_END	H	P_1	P_2	P_END
C1	VIDEO 1	150.3	183.7	267.2	434.2		634.6
C2	VIDEO 2	183.7	217.1	183.7	350.7	467.6	684.7
C3	VIDEO 3	150.3	167.0	167.0	334.0		417.5
C4	VIDEO 4	167.0	217.1	250.5	367.4	484.3	684.7
I5	VIDEO 5	100.2	150.3	150.3	267.2	384.1	617.9
I6	VIDEO 6	167.0	233.8	200.4	350.7		668
I7	VIDEO 7	217.1	250.5	250.5	400.8		851.7
I8	VIDEO 8	317.3	367.4	317.3	467.6	567.8	801.6
G9	VIDEO 9	267.2	317.3				
G10	VIDEO 10	116.9	167.0				
G11	VIDEO 11	283.9	317.3				
G12	VIDEO 12	267.2	300.6				
P13	VIDEO 13			250.5	367.4		584.5
P14	VIDEO 14			167.0	417.5		634.6
P15	VIDEO 15			66.8	283.9	367.4	551.1
P16	VIDEO 16			66.8	317.3	417.5	601.2

Table 1: Timing of each cue start in milliseconds from the onset of the stimuli video. G: gaze start G_END: gaze reaches the object; H: hand lift start; P_1:start of index finger projection; P_2:time of crossing mid section (Figure 2(a)); P_END: first point that reaches the object (i.e., not considering the hand adjustments involved to align the point even after reaching); C: congruent; I: incongruent, G: OnlyGaze; P: OnlyPoint. The lags that are being checked here is of hand movement to gaze start. Names suggested at column 1 would be used henceforth for referring to any stimuli video for simplicity

Figure 2: Still from stimuli video showing pointing directions from OnlyPoint trials

(a) (b)

4.4 Procedure

In order to maintain participants' attention, the experiment was divided into eight blocks with 50 trial videos each, leading to 400 trials for the whole experiment. At the beginning of each block, participants' gaze position was calibrated using a standard nine point calibration procedure.

Participants sat approximately 30 inch away from the display screen with an eye tracking device placed with it. Visual stimuli were projected onto the screen with 1280x720 pixel resolution.

Participants were asked to look at the fixation cross which were presented as inter-trial intervals and instructed to look at the object where they think the instructor was referring to during the stimuli videos.

4.5 Data Analysis

Eye movements were measured by an infrared light reflection recorded from the EyeTribe eye tracker with 30Hz sampling rate and spatial accuracy 0.5˚ for perfect calibration score, as declared by the manufacturer. Standard nine point calibration was performed and only "Perfect" score of five stars were accepted which indicated good spatial recording.

4.5.1 Structure of the raw data

The eye tracking data was written into text file that contains multiple time series sampled at 30 Hz.

These time series included (a) raw and smoothed *x*- and *y*-coordinates for the point of gaze on the screen separately for left and right eye and averaged between the two, (b) timestamps for each data sample at milliseconds accuracy, (c) "validity codes" for each eye indicating the reliability of tracking at each time point (state 7 = gaze coordinates of both eyes were perfectly sampled, state 4 = gaze location of one eye not recorded, state 8 = gaze locations not recorded), (d) pupil size, and (e)fixation state in True/False terms.

Provided gaze coordinates were estimated from the top-left corner of the display screen in unit of pixels. Average gaze coordinates with "fix:true" state and tracker state 7 and state 4 were used. State 8 data were discarded since provided gaze coordinates were(0,0) for both eyes which refers to either blinking or sampling error. Data loss due to sampling error in both eyes/blinks was found to be 2.32% over all the participants.

4.5.2 Fixation map

Fixation map for each condition was plotted using *x*- and *y*- coordinates of all subjects collapsed across time and corresponding kernel density estimation of *x*-coordinates was plotted.

4.5.3 Fixation percentage

Proportion of fixation was calculated by distributing the x-axis of the fixation points into three groups. Two groups accounted for the boundaries of two objects and the third group consisted the actor. The boundaries were taken to be (0-230), (240-870), (1040-1280) pixels considering the objects and actor's position in all the trial conditions and also considering the spatial error of the eye tracker (found to be ~25px[23]).

Fixation proportion for same condition but different trial videos were combined, e.g., four OnlyGaze trial videos fixation proportions were combined to give an average effect for OnlyGaze condition.

Proportion was considered for comparison since unbalanced fixations were obtained for trials and hence normalization was necessary.

4.5.4 Gaze arrival time(GAT)

Gaze Arrival Time is defined as that time at which prolonged fixation (>200 ms) was achieved at the referred object with respect to the onset of actor's gaze start. In our case, we took the last fixation on the referred object, assuming this to be the timepoint at which visual foraging by the participants were over. We ensured that the fixation duration between the next point, which was not on the object but elsewhere, was more than 200ms.

Since we are interested in knowing the integration of information provided in different time, we took the start of the gaze movement of the actor which was in lead in every videos to be the zero. Only in the case of Only Point trials, start of point movement of the actor was taken as zero.

The GATs were clubbed together for each stimuli videos for all participants. No participant based analysis was done.

31.45% of trials were unattended out of 6400 and 68.54%, 13.79% of the left trials were randomly anticipated on referred objects i.e. fixations were at the target before the start of the actor's movement.

Only 13.78% trials obtained were gaze-directed so were excluded from GAT analysis.
Again, GATs cutoff of 100-1000ms was introduced to exclude the outliers.

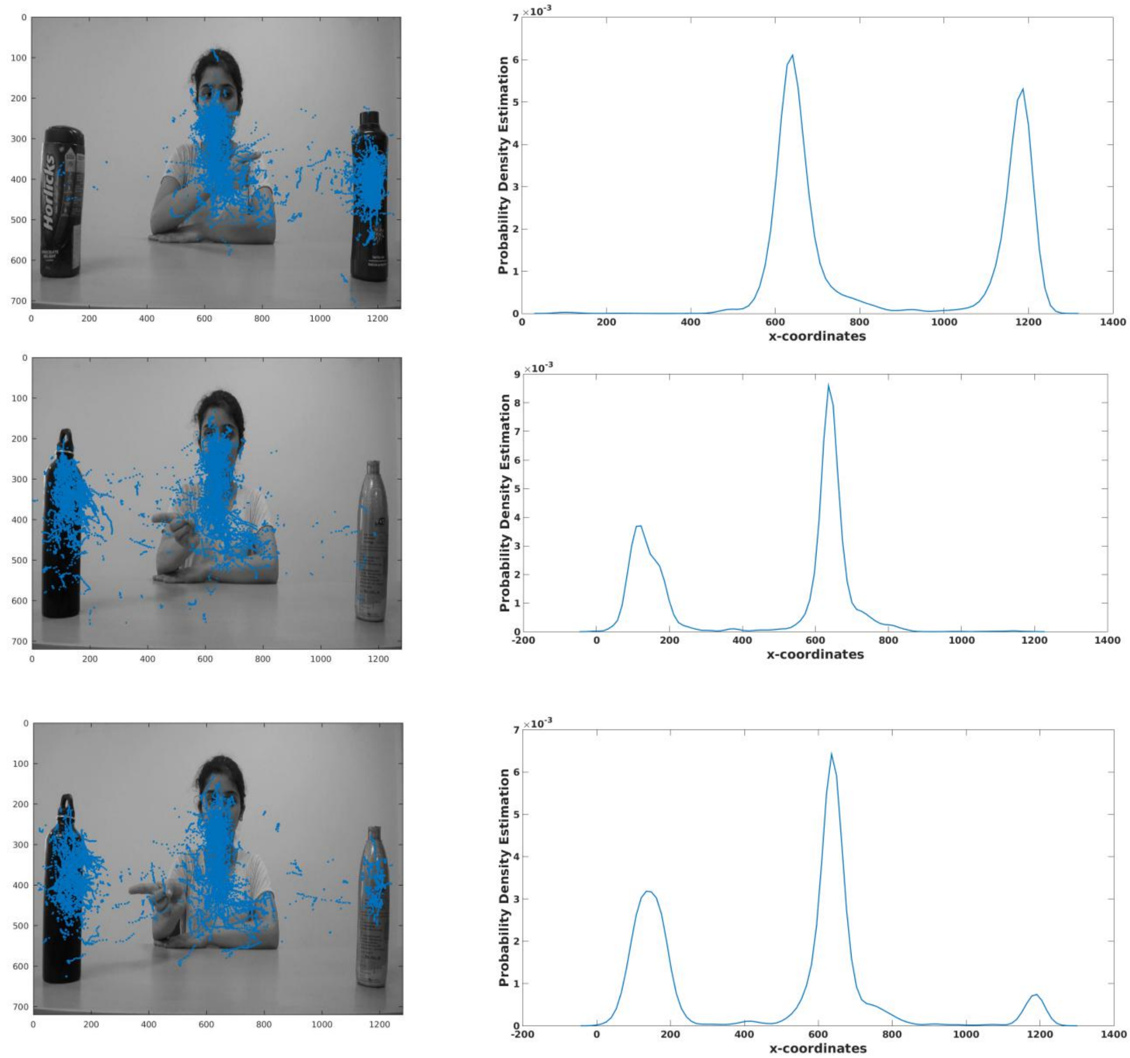
5. RESULTS

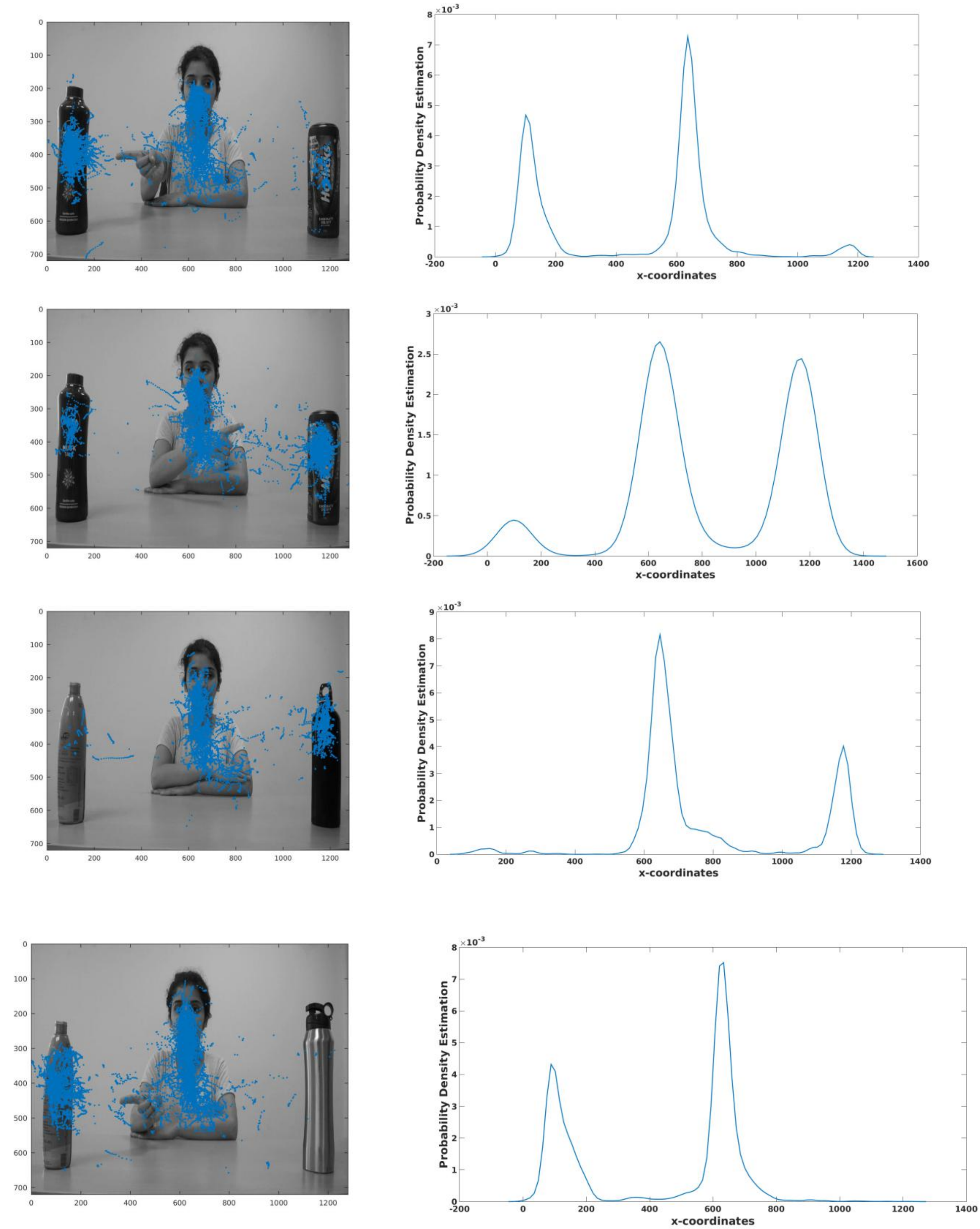
5.1 Spatial Distributions

In congruent conditions and OnlyGaze and OnlyPoint conditions, clear bimodal distributions were found which is also being depicted by the kernel density estimation plots of *x*-coordinates. Trimodal distributions were found for incongruent conditions which account for the heterogeneity of choice between gaze and point. But the third mode that is spatially representative of the gaze-directed object has probability density very less as compared to the other two clusters on actor body and the point-directed objects of incongruent trials. This was further proved by the fixation proportion across all trials against Congruent trials of 116.9 ms lag, 83.5 ms lag,16.7 ms lag, 0 ms lag, Incongruent trials of 50.1 ms lag, 33.4 ms lag, 0 ms lag, and OnlyGaze, OnlyPoint trials. Fixations were found to be maximum at the body average across all participants across conditions, ranging from 57%–61% of the total fixations while on the target objects(in case of congruent and OnlyGaze, OnlyPoint trials) and point-directed objects were fixated 25%–42%. Only 0.2%–7% fixations were on the non-targeted objects or the gaze-directed objects; 3.19%, 7.30%, 5.22% of fixations were on the gaze-directed object respectively for 50.1 ms lag, 33.4 ms lag, 0 ms lag condition of incongruent trials.

The fixation proportion were found highly significant within a condition (two-way ANOVA, $F = 201.44$, $p < 0.001$) while across conditions were not significant ($p = 1$). Thus, distributions of gaze locations were highly driven by the visual cues presented on the stimuli videos but cannot be said specifically if lags had any effect.

Figure 3: Fixation map and the corresponding kernel density distribution of x-coordinates. (a) Congruent trial with 116.9 ms lag (C1); (b) Congruent trial with zero lag (C2); (c) Incongruent trial with zero lag (I8); (d) Incongruent trial with 50.1 ms lag (I5); (e) Incongruent trial with 33.4 ms lag (I6); (f) OnlyGaze trial (G11); (g) OnlyPoint trial (P14). (read from up to down).





5.2 Gaze Arrival Time (GAT)

For comparison across conditions, ex-Gaussian probability function which consist of both normal and exponential component, was used to calculate μ, σ, τ values. Ex-Gaussian distribution is known for effective reaction time fit [29]. So μ, σ gives mean and standard deviation from normal distribution while τ gives mean for the exponential component.

Further, to test the significance among the groups wilcoxon test was performed.

	mu	sigma	tau
video1	353.6177	209.4532	129.1859
video2	333.9383	187.3838	129.5392
video3	318.4119	201.2460	152.5538
video4	314.7108	201.6525	161.3024
video5	335.8822	205.7973	149.0315
video6	411.4657	228.9302	116.6890
video7	507.3542	263.6708	44.5334
video8	350.9725	202.5017	134.9999
video9	247.0320	152.4614	129.6309
video10	252.6017	157.9508	153.2639
video11	231.4793	161.2632	142.0612
video12	222.3989	122.3377	122.2937
video13	384.4549	213.7740	120.5685
video14	421.9227	215.1286	91.5645
video15	331.1112	192.2135	139.4903
video16	349.3479	206.2515	139.5994

Table 2: Calculated parameters from ex-Gaussian distribution: mu represent the mean of the normal component while tau gives the mean of exponential component. More the tau value, more right skewed the distribution would be.

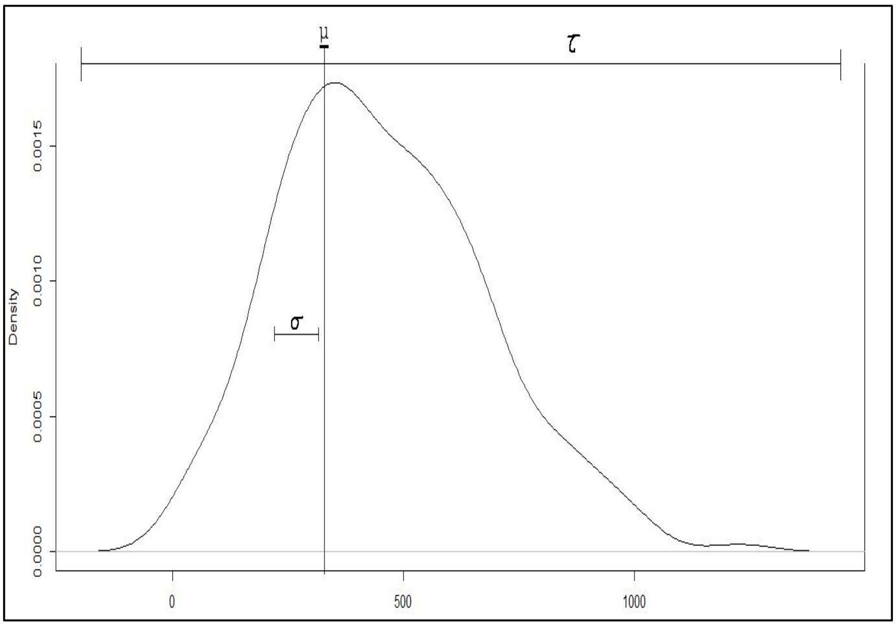


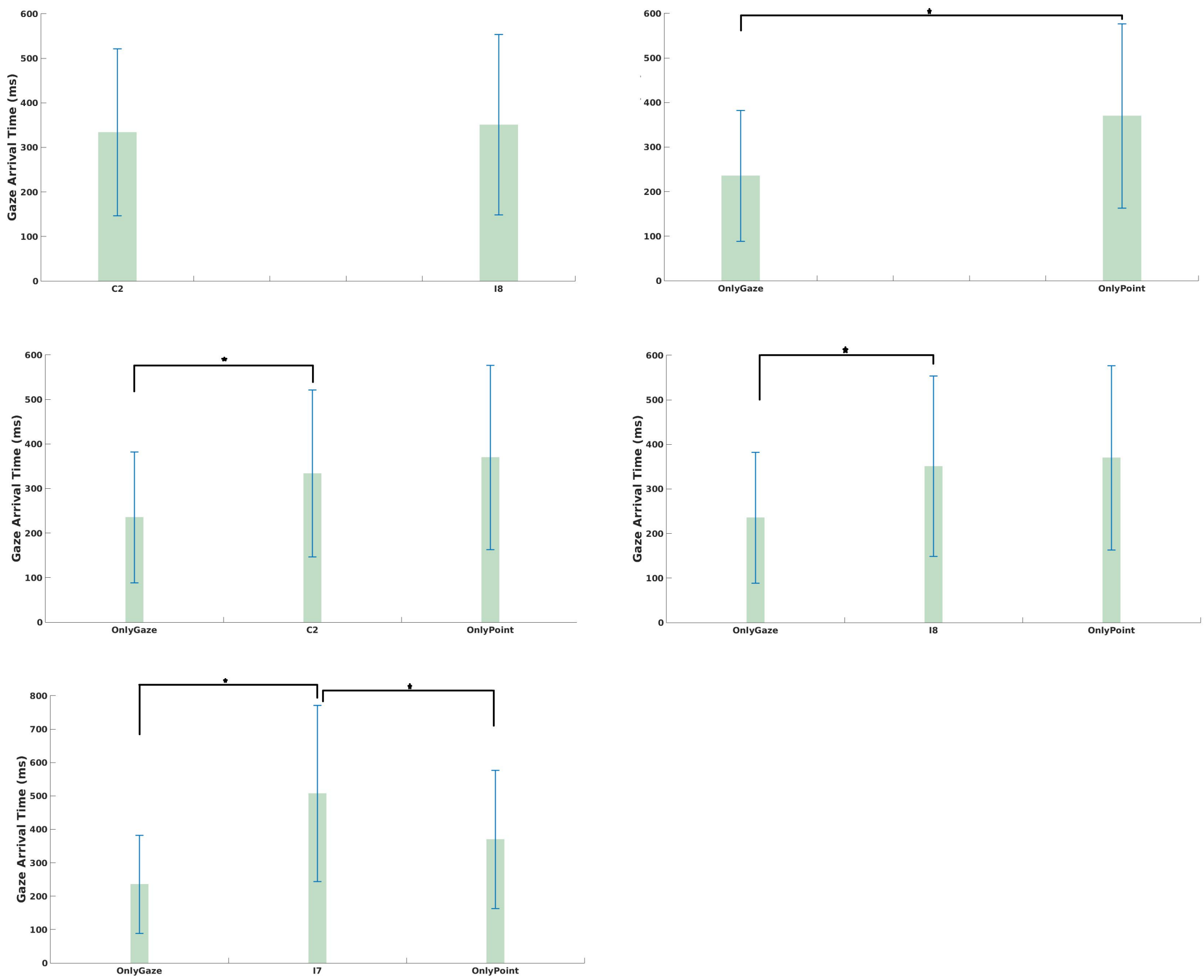
Figure 4: a simulated ex-Gaussian distribution of reaction time

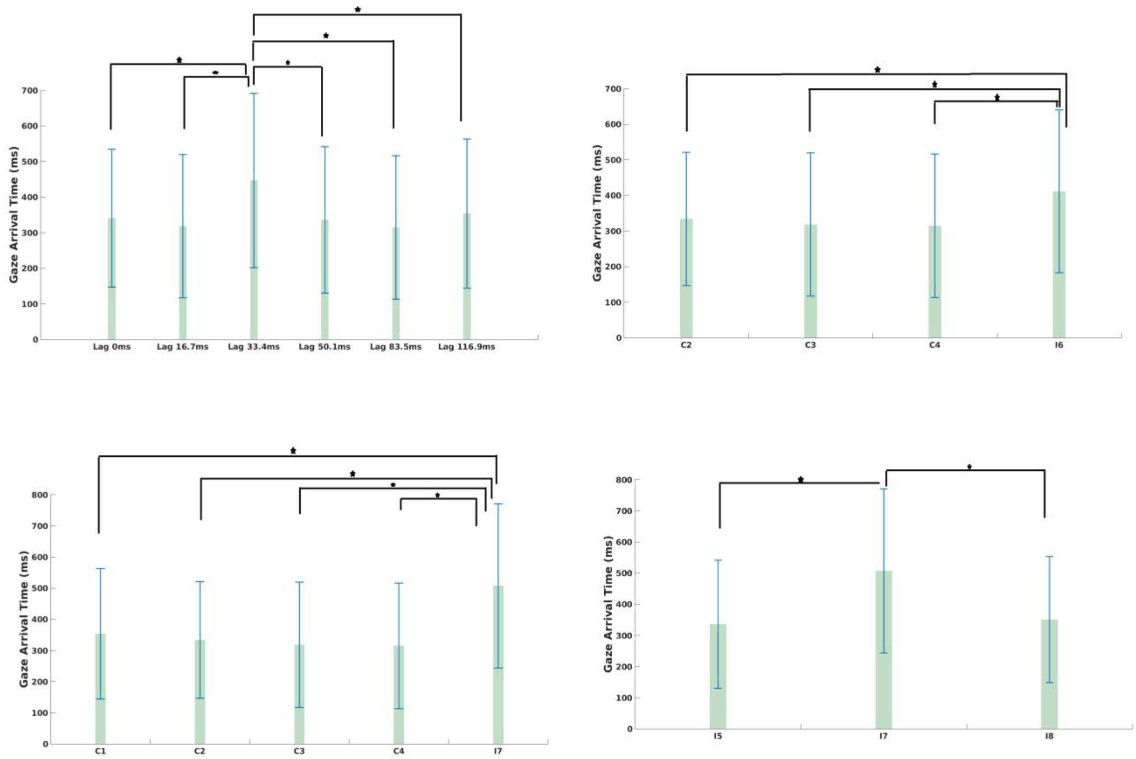
There was no significant difference found within stimuli videos of OnlyGaze and OnlyPoint trials. But GATs for OnlyPoint trials were significantly higher than OnlyGaze trials (Figure 5 (b)). OnlyGaze trials had significantly low GATs as compared to congruent and incongruent videos of zero lag but with OnlyPoint trials no significant difference was found (Figure 5 (c,d)).

Only I7 showed high GAT to OnlyPoint and OnlyGaze trials (Figure 5 (e)).The assumed main effect of congruency was checked by comparing congruent lag 0ms with corresponding incongruent lag trial (i.e. between C2 and I8). But no significance difference was found (Figure 5 (a)). But surprisingly, I7 and I6 did show significantly higher GAT to that of different congruent trials with only exception of I6 being similar to C1 (Figure 5 (g,h)).

Again, within congruent condition GAT did not differ significantly but in incongruent condition, I7 trials showed significantly higher GAT than I5 and I8 (Figure 5 (i)). With respect to lag of hand movement start to gaze start, only lag 33.4 ms showed significantly higher GAT to that of lag 0 ms, 16.7 ms, 50.1 ms, 116.9 ms.

Figure 5: GAT difference among described trials; star depicts high statistical significance with p < 0.05. read from left to right as (5a) to (5i)





6. DISCUSSIONS

Kernel density estimation plot established that majority of distributions of gaze locations were on point-directed target and actor's body in incongruent conditions. This is in line with the *Ambrosini et al. 2015* study of gaze and hand pre-shape where they showed that motor cue is more informative than gaze cue when both the information were present. Henceforth, the distributions found toward point-directed object and the actor's body in incongruent conditions is evident of the point cue being more informative or reliable source of intention conveyance.

We did not check for consistency of these fixation pattern with participants' subjective choice *per se* to validate that the gaze-directed fixations obtained were due to some participants' preference to gaze cue or was due to random fixation during visual search by the participants.

Attended trials which was measured as trials which had fixation on the referred object was >30% for all trials, except for OnlyGaze trials with only 26.56% attended. Regardless, the consistency of maximal fixation on actor was found in all conditions. One reason for maximum fixation on actor could be because, observers' were waiting for cues from the actor to predict her directional action. This may be evident if temporal fixation shift along the video length is accounted for which wasn't focused for the current analysis.

GAT analysis gave surprising results: significantly higher GAT of OnlyPoint trials to OnlyGaze trials did suggest that there was an effect of the cue presented, with gaze cue contributing to fast information conveyance than point. Consequently, congruent and incongruent trials showed significantly higher GATs but no significant difference was found as compared to OnlyPoint trials. But, incongruent stimuli video with lag 33.4 ms (I7) showed significantly increased GAT to both OnlyGaze and OnlyPoint.

Also the lag of hand movement with respect to gaze start of actor did not show expected result. It was only lag 33.4 ms stimuli video that interacted with other lags. Again, no difference within congruent videos of different lag were seen which is suggestive of lag not being a main effect for congruent stimuli. But within incongruent condition, lag 33.4 ms stimuli video (I7) showed significant increase of GAT to incongruent 50.1 ms and 0 ms lag (I5 and I8). These are all suggestive of lag not being the only effect contributing to unexpected GATs for stimuli videos of congruent and incongruent conditions. Apart from lag, it could be the actor's gaze reach time and point projection start or end time that is interacting. This may be a reason that why only incongruent lag 33.4 ms (I7 and I6) showed increased GAT to congruent videos while lag 50.1 ms and 0 ms showed similar GATs even if the conditions were incongruent or why OnlyPoint trials GATs were similar to most of the presented conditions. More post hoc analyses are required to reach any conclusion about the information integration of gaze and pointing in dynamic environment.

7. REFERENCES

1. Ambrosini, E., Pezzulo, G., & Costantini, M. (2015). The eye in hand: predicting others' behavior by integrating multiple sources of information. Journal of Neurophysiology, 113(7), 2271–2279. https://doi.org/10.1152/jn.00464.2014

2. Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. Journal of Neurophysiology, 106(3), 1437–1442. https://doi.org/10.1152/jn.00118.2011

3. Bach, P., & Schenke, K. C. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. Social and Personality Psychology Compass, 11(7), 1–17. https://doi.org/10.1111/spc3.12312

4. Conty, L., Dezecache, G., Hugueville, L., & Grezes, J. (2012). Early Binding of Gaze, Gesture, and Emotion: Neural Time Course and Correlates. Journal of Neuroscience, 32(13), 4531–4539. https://doi.org/10.1523/JNEUROSCI.5636-11.2012

5. Costantini, M., Ambrosini, E., Cardellicchio, P., & Sinigaglia, C. (2014). How your hand drives my eyes. Social Cognitive and Affective Neuroscience, 9(5), 705–711. https://doi.org/10.1093/scan/nst037

6. Elsner, C., D'Ausilio, A., Gredebäck, G., Falck-Ytter, T., & Fadiga, L. (2013). The motor cortex is causally related to predictive eye movements during action observation. Neuropsychologia

7. Flanagan, J. R., & Johansson, R. S. (2015). Action plans used in action observation, (May). https://doi.org/10.1038/nature01861

8. Friedburg, C., Allen, C.P., Mason, P.J., Lamb, T.D., 2004. Contribution of cone photoreceptors and postreceptoral mechanisms to the human photopic electroretinogram. J. Physiol. (Lond.) 556, 819–843.

9. Geiger, A., Niessen, E., Bente, G., & Vogeley, K. (2017). Eyes versus hands: How perceived stimuli influence motor actions. PLoS ONE, 12(7), 1–12. https://doi.org/10.1371/journal.pone.0180780

10. Gredebäck, G., & Falck-Ytter, T. (2015). Eye Movements During Action Observation. Perspectives on Psychological Science, 10(5), 591–598. https://doi.org/10.1177/1745691615589103

11. Grosbras, M. H., Laird, A. R., & Paus, T. (2005). Cortical regions involved in eye movements, shifts of attention, and gaze perception. Human Brain Mapping, 25(1), 140–154. https://doi.org/10.1002/hbm.20145
12. Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior, 9(4). https://doi.org/10.1016/j.tics.2005.02.009
13. Holmqvist, Kenneth. Eye Tracking: A comprehensive guide to methods and measures (pp. 22-23). OUP Oxford. Kindle Edition.
14. König, P., Wilming, N., Kietzmann, T. C., Ossandon, J. P., Onat, S., Ehinger, B., … Kaspar, K. (2016). Eye movements as a window to cognitive processes. Journal of Eye Movement Research, 9(5), 1–16. https://doi.org/10.16910/jemr.9.5.3
15. Land, M. F. (2006). Eye movements and the control of actions in everyday life, 25, 296–324. https://doi.org/10.1016/j.preteyeres.2006.01.002
16. Land, M.F., 1999. Motion and vision: why animals move their eyes. J. Comp. Physiol. A 185, 341–352
17. Materna, S., Dicke, P. W., & Thier, P. (2008). The posterior superior temporal sulcus is involved in social communication not specific for the eyes. Neuropsychologia, 46(11), 2759–2765. https://doi.org/10.1016/j.neuropsychologia.2008.05.016
18. Mennie, N., Hayhoe, M., & Sullivan, B. (2007). Look-ahead fixations: Anticipatory eye movements in natural tasks. Experimental Brain Research, 179(3), 427–442. https://doi.org/10.1007/s00221-006-0804-0
19. Neuroscience. 2nd edition.Purves D, Augustine GJ, Fitzpatrick D, et al., editors.Sunderland(MA):SinauerAssociates;2001.https://www.ncbi.nlm.nih.gov/books/NBK11156/
20. Peeters, D., Snijders, T. M., Hagoort, P., & Özyürek, A. (2017). Linking language to the visual world: Neural correlates of comprehending verbal reference to objects through pointing and visual cues. Neuropsychologia, 95(September 2016), 21–29. https://doi.org/10.1016/j.neuropsychologia.2016.12.004
21. Pierno, A. C., Tubaldi, F., Turella, L., Grossi, P., Barachino, L., Gallo, P., & Castiello,U.(2009). Neurofunctional modulation of brain regions by the observation of pointing and grasping actions. Cerebral Cortex, 19(2), 367–374. https://doi.org/10.1093/cercor/bhn089

22. Pierno AC, Becchio C, Tubaldi F, Turella L, and Castiello U. Motor ontology in representing gaze– object relations. Neurosci lett 430(3): 246-251, 2008
23. Popelka, S., Stachoň, Z., Šašinka, Č., & Doležalová, J. (2016). EyeTribe tracker data accuracy evaluation and its interconnection with hypothesis software for cartographic purposes. Computational Intelligence and Neuroscience, 2016. https://doi.org/10.1155/2016/9172506
24. Sato, W., Kochiyama, T., Uono, S., & Yoshikawa, S. (2009). Commonalities in the neural mechanisms underlying automatic attentional shifts by gaze, gestures, and symbols. NeuroImage, 45(3), 984–992. https://doi.org/10.1016/j.neuroimage.2008.12.052
25. Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. Proceedings of the Symposium on Eye Tracking Research & Applications - ETRA '00, (October 2014), 71–78. https://doi.org/10.1145/355017.355028
26. Schmitow, C., Kochukhova, O., & Nyström, P. (2016). Social perception: How do 6-month-old infants look at pointing gestures? Infant Behavior and Development, 42, 152–156. https://doi.org/10.1016/j.infbeh.2015.12.001
27. Science, P., & Gredeb, G. (2015). Eye Movements During Action Observation, (September). https://doi.org/10.1177/1745691615589103
28. Shapovalova, N., Raptis, M., Sigal, L., & Mori, G. (2013). Action is in the Eye of the Beholder: Eye-gaze Driven Model for Spatio-Temporal Action Localization. Advances in Neural ….
29. Whelan, R. (2008). Effective analysis of reaction time data. The Psychological Record, 58, 475–482. https://doi.org/10.1007/BF03395630
30. https://github.com/EyeTribe/documentation
31. http://theeyetribe.com/dev.theeyetribe.com/dev.theeyetribe.com/general/index.html#intro
32. http://www.pygaze.org/2015/06/eyetribe/