LargeEyeTrackingData

Monday, October 27, 2025 11:25 PM

This project analyzes eye-tracking data from a visual search and free-viewing experiment.

Paper: A very large eye tracking dataset of free viewing and visual search
Original data and analysis code download from: https://osf.io/kf4sb/files/nv8cu

Intro to dataset

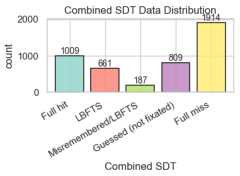
All data organized by task: `search/` and `freeviewing/`.
Free viewing (1,248 ppl) and searching(2,827 ppl) are completely **independent** samples of participants (zero overlapped participant).
This dataset comes with its own analysis code, which **only** detects fixations without saccades.
For more details of the code and data, see README in **Github**:
https://github.com/wppqywq/large_gaze_data.

- raw_data/
  - Gaze samples: position (x, y), timestamp, pupil size, validity flag
  - Event markers: task transitions, screen changes, response events.

  -> Detected fixation points computed using I2MC algorithm.
  Only includes fixations with duration >= 60ms within first 10 seconds of task.

- intermediate_data/
  Aggregated data and participant metadata. Contains:
  - participant_info.csv:
    [ID, Gender, Age, Task, Image, Target & Shape,
    Response, Response outcome(check shape=response, Hit/False Alarm/Nothing found),
    Valid demographics & task]
  - compiled_gaze.csv: all valid gaze samples concatenated (first 10s only)
    [x, y, time_from_start, ID]
    -> I2MC
  - compiled_fixations.csv: all valid fixations concatenated, columns same as `/fixation_data/`

  - for search task, also `accuracy_with_gaze.csv`:
    including basic [ID, Target location(str & px), Response outcome(Hit/False Alarm/Nothing found)],
    also joint responce and gaze data together:
    - Response hit: whether the response is correct or not
    - Gaze hit: whether eye movement saw the target (within 1.5 degrees and lasting ≥0.2 seconds)
    - Combined hit: Response is correct AND eye movement is seen
    - Combined SDT:
      - Full hit: Response correct + Detection correct
      - Guess: detection incorrect + response correct
      - Full miss: No detection + response Incorrect
      - LBFTS (Looked But Failed To See): response incorrect + detection correct (Response outcome = 'Nothing found')
      - Misremembered/LBFTS: same as LBFTS, but Response = 'False Alarm'



    - Gaze hit onset/duration: first gaze at the target, start time and duration
    - Refixation count: count that the target is refixed after seeing it (0 = only seen once).
    - Nearest gaze: (visual angle) between the nearest gaze point and the target.

Experiment Configs

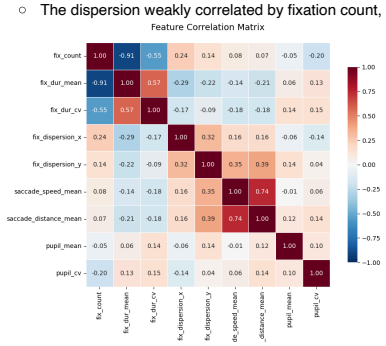
- Eye-tracker sampling rate: 60 Hz
- Minimum fixation duration: 60 ms
- Merge distance for close fixations: ~1 degree visual angle (≈ 44.8 px)
- Target counted as “fixated” if within 1.5 degrees for at least 0.2 s
- Fixation detection: I2MC algorithm (only fixations ≥ 60 ms within first 10 s)

Free Viewing dataset

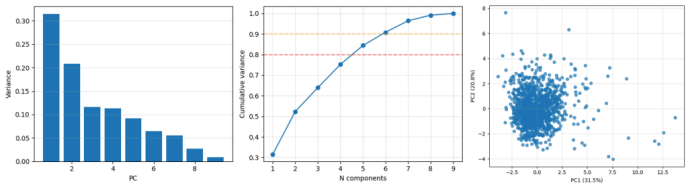
Above is a quick analyze on the free viewing dataset. In this stage, we consider the target searching dataset providing more information and mainly focus on that.

We performed PCA with 10 Features:
['fix_count', 'fix_dur_mean', 'fix_dur_cv', 'fix_dispersion_x', 'fix_dispersion_y', 'saccade_speed_mean', 'saccade_distance_mean', 'pupil_mean', 'pupil_cv']

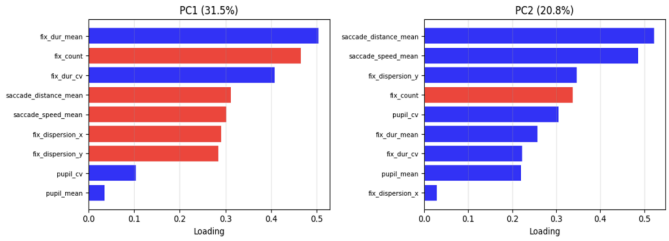
- The correlation between features are quite intuitive:
  - Three groups: (i) fixation params strongly correlated (-> redundant); (ii)saccade and visual angle dispersion range correlated; (iii)pupil variables isolated.
  - The dispersion weakly correlated by fixation count, more fixation wider dispersion.



- PCA analysis
  - PC1 explains ~30% of total variance. PC2 explains ~20%. 6 PCs explains 90%.



- Loading plot
  - PC1 has high loadings on fixation behavior; PC2 mainly dominants by Saccade Dynamics



Target Search Dataset

Features:

Total **16** features were used for correlation, PCA, and clustering:

- Fixation basic:
  `fix_count, fix_dur_mean, fix_dur_cv,`
  `fix_dispersion_x, fix_dispersion_y,`

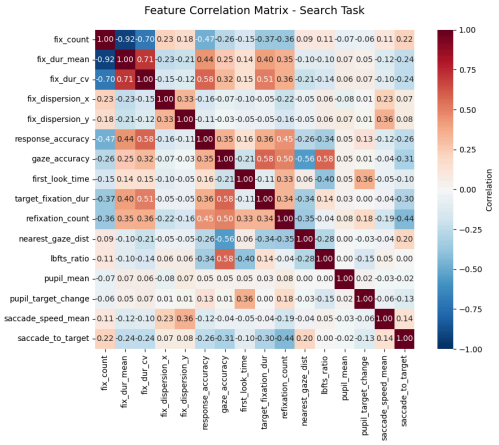
- `Binary params:`
  `response_accuracy, gaze_accuracy, (1=hit, 0=incorrect)`
  `lbfts_ratio,`

- `If gaze hit taget:`
  `refixation_count, target_fixation_dur, (=median if not)`
  `first_look_time, nearest_gaze_dist,(=median if not)`

- `Pupil and saccade:`
  `pupil_mean, saccade_speed_mean,`
  `pupil_target_change (mean pupil in a ±0.1 s window around first_look_time minus pupil_mean.) (=median if no gaze hit)`
  `saccade_to_target (cosine similarity between the saccade leading to the fixation nearest first_look_time and the vector from that fixation to the target location) (=median if no gaze hit)`

Correlations

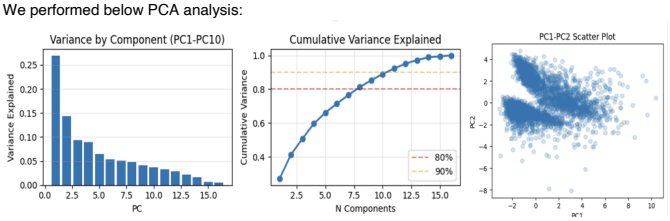
With these features, we plot the correlation matrix:



- Like free viewing, Fixation measures are highly correlated. Fixation count and duration almost negatively proportional.
- Response accuracy align with gaze accuracy at only 0.35.
- Accuracy-related measures form a connected block. Gaze accuracy related to the fixation duration (no matter target fixation or all fixations). Higher fixation duration and variation increase both accuracy, indicating high attention level.
- Refixation also increase accuracy. And refixation count has negative relation with fixation count!
- LBFTS negatively to first_look_time, delayed may increase unawareness.
- Pupil change at target positively relate to first fixation time;
  saccade accuracy to target (large if hit target) negatively related to the refixation count.

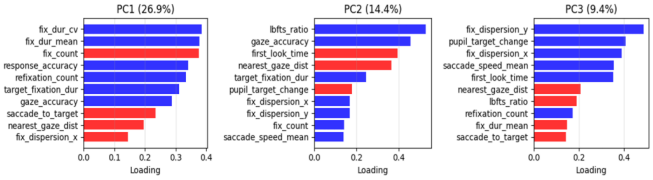
PCA analysis

We performed below PCA analysis:



- PC1 explains 26.9% of total variance. PC2 explains 14.4%, and PC3 explains 9.4%. The first 10 PCs can rebuild ~90% infomations.
- The PC1–PC2 scatter plot (right) reveals a clear multimodal structure, motivating the later GMM clustering.

Component Loadings and Accuracy corr:



	response_accuracy	gaze_accuracy	lbfts_ratio
PC1	0.71	0.60	–0.04
PC2	–0.16	0.69	0.80
PC3	0.18	0.01	–0.23

- PC1: Fixation stability and overall accuracy
  High positive loadings: fix_dur_mean, fix_dur_cv, response_accuracy, refixation_count, target_fixation_dur, gaze_accuracy
  Negative loadings: fix_count, nearest_gaze_dist, saccade_to_target, fix_dispersion_x
  PC1 captures overall performance, higher response and gaze accuracy..
  Participants with higher scores tend to make fewer but longer fixations, sustain attention on the target, and perform with higher visual and response accuracy.

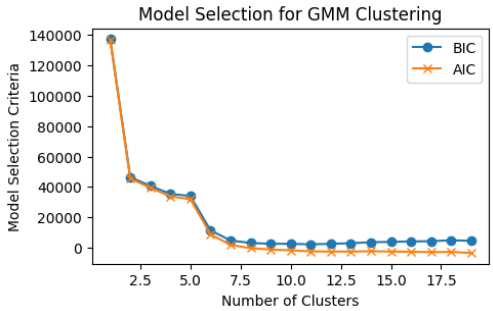
- PC2: Awareness and detection dissociation
  - Positive: lbfts_ratio, gaze_accuracy, first_look_time
  - Negative: nearest_gaze_dist, target_fixation_dur, pupil_target_change

  Higher scores indicate participants who often looked at the target (high gaze_accuracy) but did not consciously report it (high lbfts_ratio). -> seeing but not knowing.

- PC3: Spatial and pupil variability

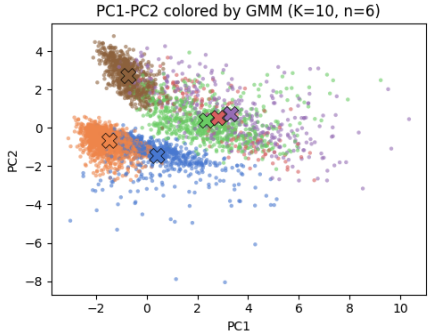
GMM Clustering

Gaussian Mixture Models (GMMs) were fit with cluster numbers ranging from 2 to 20. We use first 10 PCs, and evaluate the model with both AIC and BIC criteria.
Both curves decreased sharply up to around 6 clusters as the plot below:



So we choose best_n=6 for clustering, result in a cluster size distribution:
{1: 2073, 2: 706, 5: 654, 0: 652, 4: 360, 3: 137} for total 4583 participants.
The uneven distribution suggests one large, dominant behavioral group (Cluster 1), and several smaller subgroups with distinct visual search styles.

The resulting PC1–PC2 plot shows six clearly separable groups, distributed mainly along the two principal axes representing fixation–accuracy (PC1) and awareness–detection (PC2).

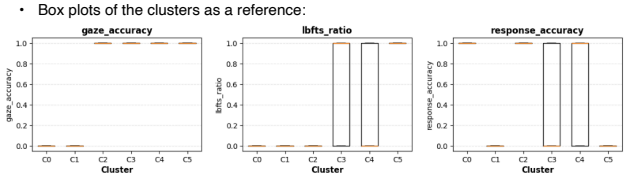


Cluster Characteristics:

- Cluster 0: n=652 (14.2%) response_accuracy(+1.08) | gaze_accuracy(-0.83) | fix_dur_cv(+0.62) | nearest_gaze_dist(+0.62) | fix_count(-0.54) ...
  -> Guesser
  High response accuracy but low gaze accuracy suggests many correct reports without actually fixating on the target.
  They show short target fixations, low fixation count, and high fixation duration variability, meaning they make fewer but unstable eye movements.
  The large nearest gaze distance implies their gaze often stayed away from the target area.

- Cluster 1: n=2073 (45.2%) gaze_accuracy(-0.83) | response_accuracy(-0.72) | fix_dur_cv(-0.54) | target_fixation_dur(-0.48) | lbfts_ratio(-0.48) ...
  -> Full miss
  - Largest and most average group.
  - Both response and gaze accuracy are low, fixation duration short, refixations rare, and search dispersion wide.

- Box plots of the clusters as a reference:



- many short fixations and high gaze distance, reflecting a fast but shallow scanning style without consistent engagement.

- Cluster 2: n=706 (15.4%) response_accuracy(+1.39) | gaze_accuracy(+1.21) | target_fixation_dur(+1.05) | fix_dur_cv(+0.94) | refixation_count(+0.80) ...
  -> Full hit
  High in both accuracies, long target fixations, and many refixations with low nearest gaze distance. Making fewer fixations, longer, stable, and well-directed.

- Cluster 3: n=137 (3.0%) refixation_count(+2.08) | gaze_accuracy(+1.21) | fix_dispersion_x(-1.03) | lbfts_ratio(+1.01) | fix_count(-1.01) …
  -> Slow action
  Very high refixation count and long fixations, but late first look to the target.
  They eventually find it, as gaze accuracy and target fixation duration are high, but search dispersion is narrow, and might not be able to answer correctly.

- Cluster 4: n=360 (7.9%) saccade_to_target(-2.73) | refixation_count(+1.39) | gaze_accuracy(+1.21) | target_fixation_dur(+1.12) | fix_dur_mean(+0.86) …
  -> non-linear people
  high response and gaze accuracy, long fixations, many refixations, but very low saccade-to-target, maybe indicating complex search paths.
  Why are they not making correct response?

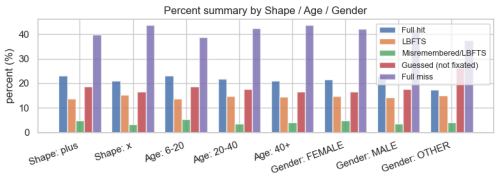
- Cluster 5: n=654 (14.3%) lbfts_ratio(+2.10) | gaze_accuracy(+1.21) | first_look_time(-1.11) | response_accuracy(-0.72) | nearest_gaze_dist(-0.57) Cluster …
  -> LBFTS

  have early first looks but short stay times, low fixation variability, and low gaze distance, showing that they reach the target quickly but leave before encoding it.
  Early detection might make people forgetting.

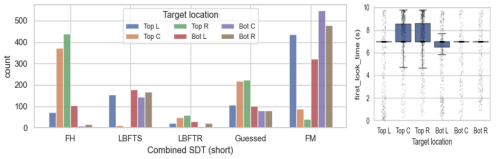
Demographic and Stimulus Effects

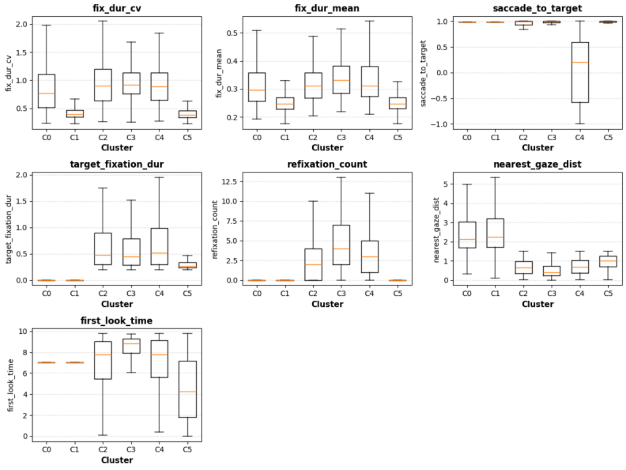
To verify that the clustering results were not confounded by demographic or stimulus factors, we examined awareness outcomes across shape, age, gender, and target position.

- Percent summaries showed highly similar distributions across all groups. Younger participants tend to get higher accuracy. Shape 'x' is slightly more difficult to find.



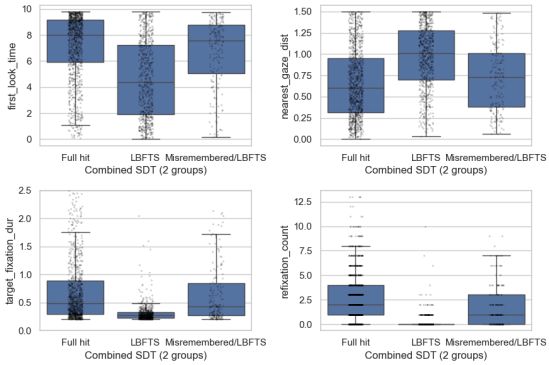
- Target position introduced some effects. Full hits were more frequent for top-center and top-left (and first-look times were later), whereas full miss is more often in top-right and lower regions.





Some box plot with different participants groups

- Study when Gaze hit = true (Full hit / LBFTS / Misremembered-LBFTS)



  - LBFTS with 'Nothing found': an eariler initial contact, larger nearest_distance (unlikely to look back), shorter fixation duration at target, and less refixation count. These behavior is close to those who truly didn't see the subjects.
  - LBFTS with misremembered is closer to full hit.
- By Combined SDT: Full miss and LBFTS with 'nothing found' show higher fixation counts (lower duration mean) and lower coefficient of variation than Full hit.
- Misremembered-LBFTS & Guessed is very similar, in between of full hit and full miss. This indicating that, the performance of seeing something but not knowing it and not seeing it but guessing it may be somewhat similar.