

Eye-Tracking Reading Experiment Fixation Analysis Using R

1. Project Background & Dataset Introduction

This case study analyzes eye-tracking data collected during a reading experiment. The goal of the experiment is to investigate reader fixation behaviors across different sections of reading material, referred to as Areas of Interest (AOIs). The dataset comprises 10 sheets, each representing a unique reading material. Every sheet has 20 columns, indicating fixation data for 20 AOIs (from s01 to s20), which correspond to different parts of the reading material.

Each dataset helps examine how long participants fixate on specific AOIs while reading different sections. This study uses R to perform detailed analysis and generate visual insights into the fixation patterns.

2. Analysis Objectives

Each of the following steps reflects the specific objectives of this case study:

1. Analyze Fixation Frequency across all AOIs, sorted by fixation count.
2. Create a Heatmap to visualize fixation patterns across AOIs and datasets.
3. Compare Fixation Count Between Datasets for each AOI.
4. Examine the Distribution of Total Fixations per Participant across datasets.
5. Plot a Box Plot to understand fixation count variability across participants.
6. Analyze Fixation Trends across datasets, splitting AOIs into two groups (s01 to s10 and s11 to s20).

3. Analysis Process

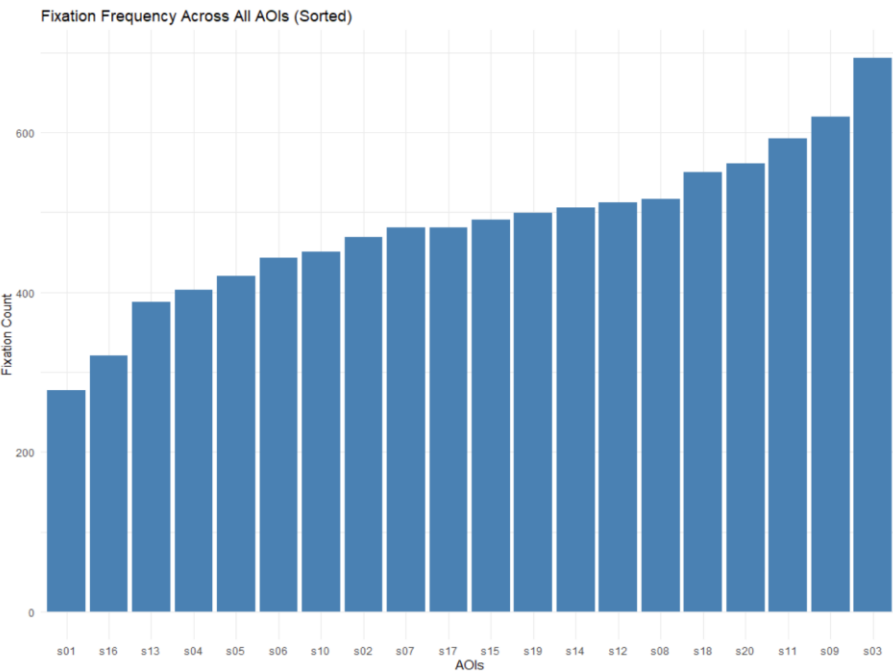
The data was read and merged for analysis using R. Missing values were replaced with 0, and datasets were reordered to ensure proper visualization.

The analysis and visualization were done using ggplot2 for creating graphs. I used data manipulation functions from the dplyr package to clean and prepare the data.

Bar charts, heatmaps, line charts, box plots, and histograms were selected to explore the fixation data from multiple perspectives.

4. Insights from the Analysis

4.1 Varying Levels of Reader Engagement



This bar chart presents an overview of the fixation counts for twenty different Areas of Interest (AOIs) from a reading experiment. The x-axis lists these AOIs, from s01 to s20, while the y-axis shows the total number of fixations, which range from just over 250 to more than 650. The AOIs are arranged in ascending order based on the fixation count, providing a clear picture of how attention was distributed across different sections of the reading material.

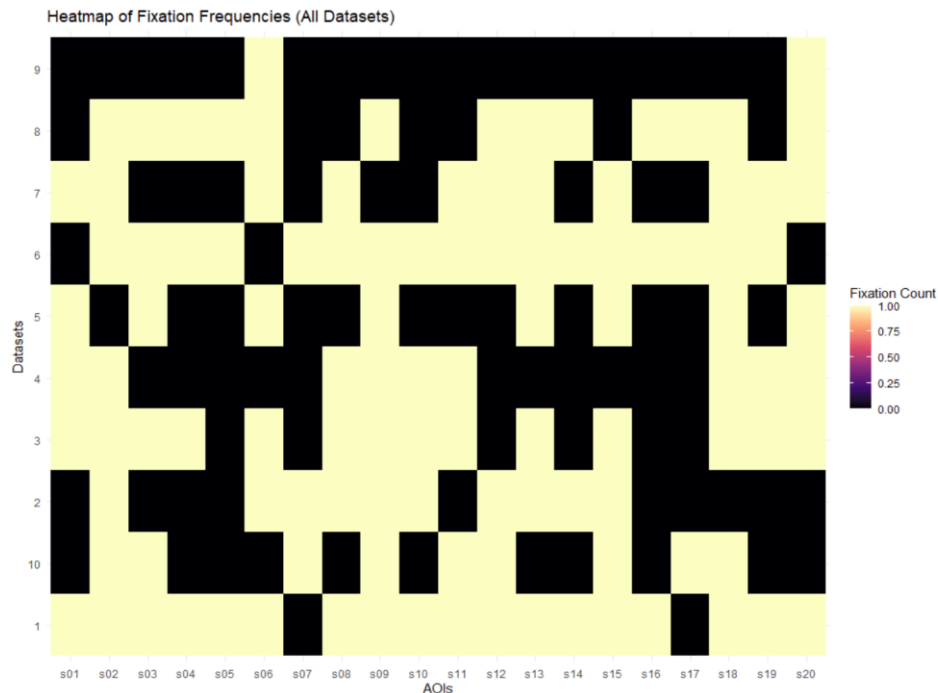
At the lower end of the spectrum, AOIs such as s01, s16, and s13 display the fewest fixations, with s01 registering just above 250 fixations. This suggests that these sections of the reading material were less engaging to participants, or perhaps they required less cognitive focus. Following them, AOIs like s04 and s05 show slightly higher fixation counts, approaching 400, but they too remain on the lower side of the attention spectrum.

In the midrange, AOIs like s06, s10, and s02 demonstrate moderate levels of engagement, with fixation counts hovering between 450 and 500. These sections appeared to hold participants' attention to a moderate extent but did not stand out as particularly engaging or attention demanding.

At the upper end of the scale, we observe significantly higher fixation counts for AOIs s03, s09, and s11. These AOIs stand out, with s03 receiving the highest number of fixations, exceeding 650. The high fixation count in these sections suggests that participants focused more intensely on these parts of the reading material, possibly indicating that these AOIs contained more complex or engaging content.

Overall, the trend shows a steady increase in fixation counts as we move from s01 to s03, with some AOIs demanding significantly more attention than others.

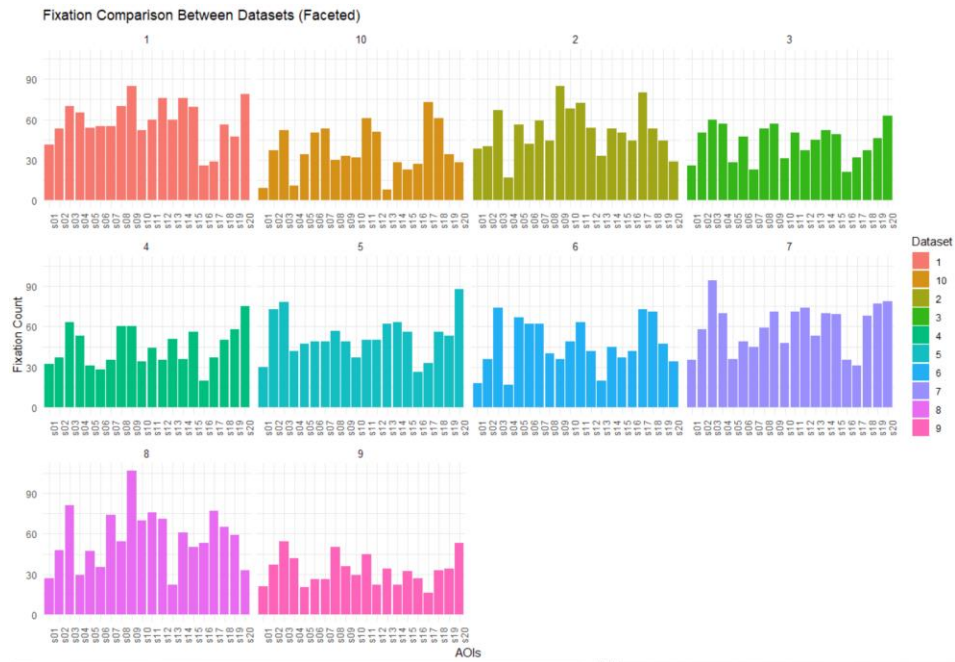
4.2 Uneven Fixation Distribution Across AOIs



This heatmap visualizes the distribution of fixation counts across different Areas of Interest (AOIs) and datasets in the reading experiment. The x-axis represents the AOIs, ranging from s01 to s20, and the y-axis represents the ten datasets involved in the analysis. The color gradient on the right indicates the fixation count, with values ranging from 0 (black) to 1 (yellow), representing the presence or absence of fixations across the AOIs for each dataset.

AOIs such as s03, s11, and s09 exhibit consistent engagement across most datasets, as indicated by the yellow shading. In contrast, AOIs like s01, s05, s16, and s17 show much less engagement, with more black regions, signifying low or no fixation activity. This variation suggests that different datasets elicited different levels of attention across the AOIs, which may be attributed to differences in the reading material or participant focus.

4.3 AOI Engagement Differs by Dataset

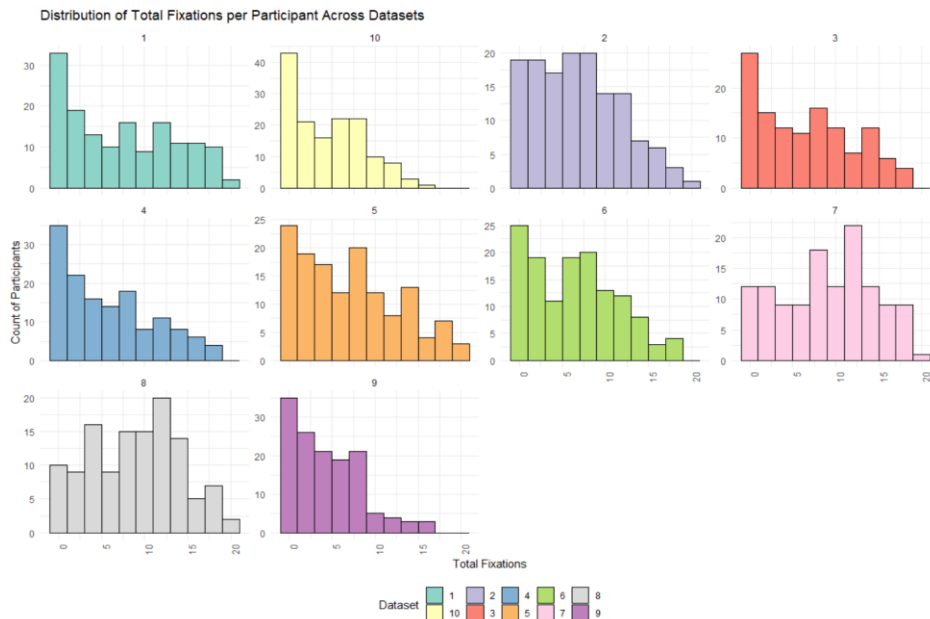


This faceted chart presents the fixation count across Areas of Interest (AOIs) for each dataset involved in the reading experiment. Each facet (or small individual graph) represents a dataset, with the x-axis showing the AOIs (s01 to s20) and the y-axis displaying the fixation count.

Key observations from this chart reveal that some AOIs, such as s03, s09, s10, and s20, consistently draw higher fixation counts across multiple datasets, indicating that these sections of the reading material were more engaging for participants. Conversely, AOIs like s01, s05, and s17 tend to show lower fixation counts across most datasets, suggesting these parts received less attention during the reading tasks.

From another perspective, we can find out that Dataset 1 shows a clear peak in AOI s03 with fixation counts nearing 90, while Dataset 10 has higher engagement distributed across AOIs s09, s10, and s20.

4.4 Low Fixation Counts Dominate Participants



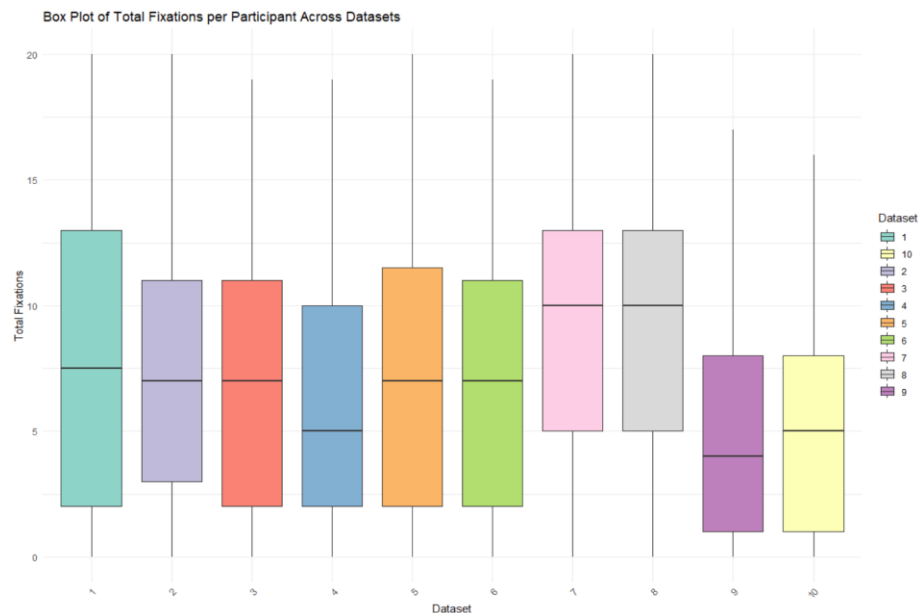
This faceted chart illustrates the distribution of participants' total fixations across various datasets. Each individual chart represents one dataset, with the x-axis showing the total number of fixations and the y-axis showing the count of participants who had that number of fixations.

This chart shows that most participants had relatively low fixation counts across all datasets, with a sharp decline in participant numbers as the fixation count increases. For instance, in Dataset 1, the highest count of participants (more than 30) had fewer than 5 fixations. This trend is echoed across the other datasets, particularly in Datasets 2, 4, 6, and 9, where the majority of participants also had fewer fixations.

However, some datasets display slight variations. For example, Dataset 10 shows a somewhat broader distribution of fixation counts, with a significant number of participants exhibiting between 5 and 10 fixations. Additionally, Dataset 7 presents a more evenly spread distribution compared to others, where participants had fixations between 5 and 15.

In summary, most participants tend to have low fixation counts, and the spread of these fixations varies depending on the dataset.

4.5 Dataset 1 Shows Highest Total Fixations



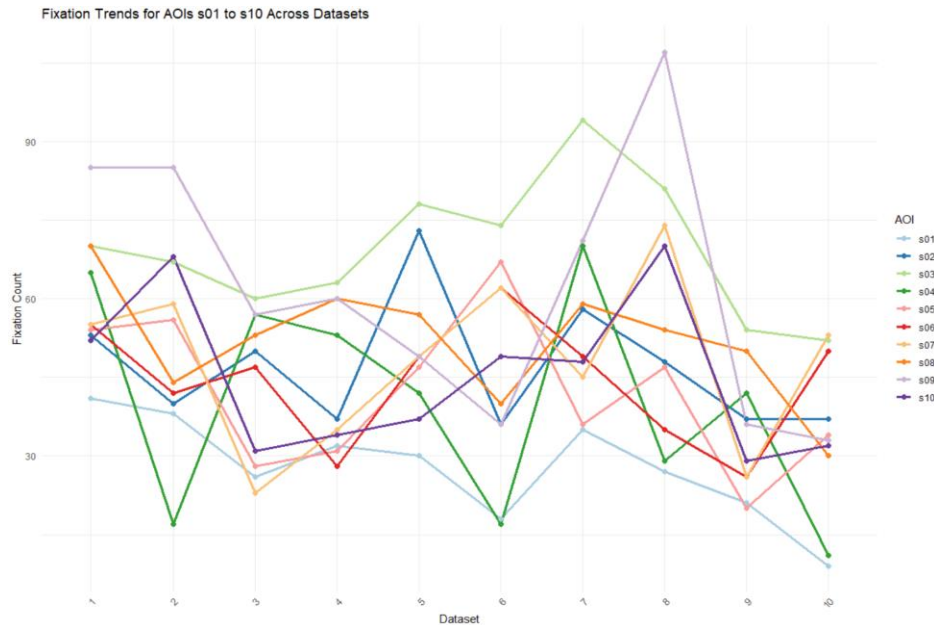
This box plot illustrates the distribution of total fixations per participant across ten different datasets. Each colored box represents a dataset, with the central line in each box signifying the median number of fixations for participants. The top and bottom edges of the box represent the interquartile range (IQR), showcasing the middle 50% of fixation counts. The whiskers extending from the boxes indicate the range of fixation counts, excluding outliers, which are represented by individual points beyond the whiskers.

Interestingly, Dataset 1 displays the highest median fixation count compared to other datasets, with values clustering around the 10-15 fixation mark. Datasets such as 4, 5, and 6 have a broader IQR, indicating greater variability in participants' fixation counts, while Dataset 9 shows one of the smallest ranges, suggesting that participants' fixation counts are more uniform in this dataset.

An outlier pattern can be noticed in several datasets, including 1, 5, and 7, where a small number of participants demonstrated either particularly high or low fixation counts compared to the rest. Dataset 9 also stands out due to its relatively narrow range of fixation values, where most participants clustered around a median value of 10 fixations.

In a nutshell, Dataset 1 consistently shows a higher fixation engagement, while Datasets 9 and 6 reveal tighter distributions around their medians.

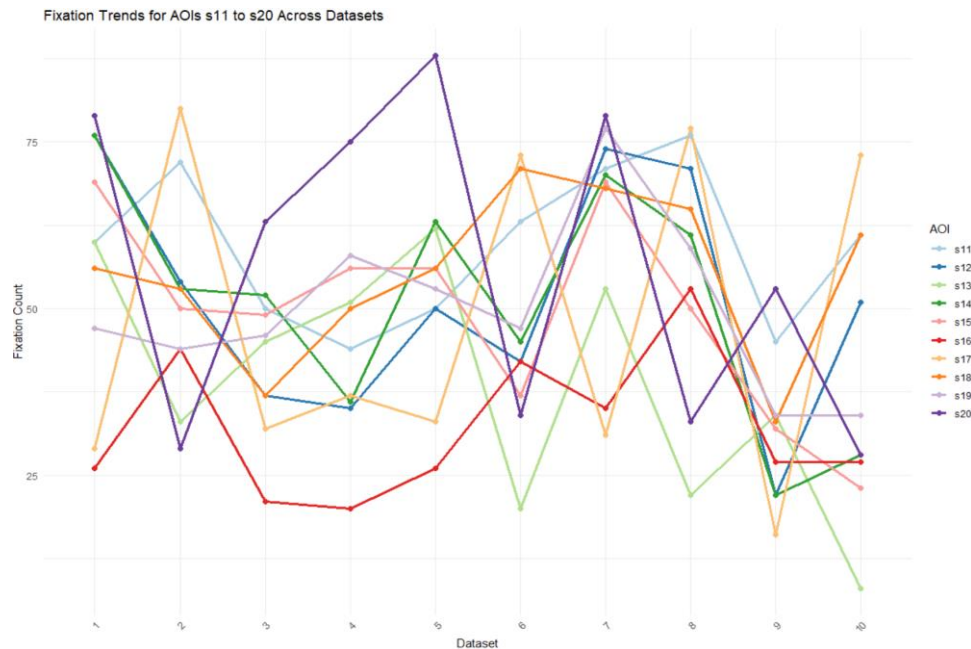
4.6 AOIs s10 and s07 Dominate Focus



This line chart displays fixation trends for AOIs (Areas of Interest) from s01 to s10 across multiple datasets, highlighting how fixation counts fluctuate for each AOI. A few noticeable peaks can be observed, such as AOI s10 having consistently high fixation counts, especially in datasets 6, 9, and 10. Similarly, AOI s07 also shows an upward trend across several datasets.

Conversely, AOIs like s02 and s06 demonstrate more variability, with fixation counts dropping significantly in certain datasets. AOIs such as s04 and s05 exhibit relatively stable, moderate fixation counts without dramatic fluctuations across datasets.

4.7 Fluctuating Fixation Patterns Across AOIs



In this line chart, I observe the fixation trends for AOIs s11 to s20 across different datasets. AOI s20 and s11 show significant variability, with s20 reaching its peak fixation count at over 75 in some datasets, while others show much lower fixation counts. Other AOIs like s17 and s12 also show fluctuating trends, though none surpass s20 in maximum fixation. There is no consistent upward or downward trend across datasets, indicating that the focus on specific areas of interest (AOIs) differs greatly depending on the dataset, with peaks and troughs distributed irregularly. The wide variance suggests that attention may be more dispersed among these AOIs than in others.

5. Conclusion

This analysis provided an in-depth look into how participants engaged with different sections of reading material based on eye-tracking fixation data. By examining fixation trends, frequencies, and distributions, I identified several key insights:

- 1) Certain AOIs (s03, s09, s12) consistently attracted higher attention, while others like s01 and s16 are less engaging.
- 2) Dataset variability: Some reading materials (such as dataset 10) were consistently more engaging across a broader range of AOIs, indicating potential differences in content structure or presentation.

The results provide actionable insights for improving reading material design and understanding reader behavior, which could inform the creation of more engaging and attention-capturing content.