Orthographic Processing in School-age Children - Descriptive Statistics Eye Tracking Data

Brief Introduction

Orthographic awareness (OA)—the ability to recognize and evaluate letter patterns based on the rules and regularities of a writing system—is an important component of reading development. In alphabetic languages such as English, OA supports both decoding and word recognition by enabling readers to distinguish between more and less probable letter sequences.

The current study investigates orthographic processing in school-age children (grades 3–6) across three groups:

Dyslexia (n = 28)

Compensated dyslexia (n = 17)

Typical development (TD) (n = 23)

By examining how these groups differ in their orthographic decision-making, the study aims to better understand the mechanisms underlying persistent and resolved reading difficulties.

Method

Participants were 68 school-age children in grades 3 through 6, divided into three groups based on their reading profile: 28 children with diagnosed dyslexia, 17 children with compensated dyslexia who had overcome earlier reading difficulties, and 23 typically developing readers.

Each child completed an 18-trial visual-world eye-tracking task programmed in E-Prime. On each trial, four non-word letter strings appeared on the screen — one high-probability ("high-ortho") string that closely followed English orthographic patterns; one low-probability ("low-ortho") string with legal but less common letter sequences; one string that violated English orthographic rules ("illegal"); and one unpronounceable string. The question prompt at the top of the screen read: "Which word looks most like a real English word?" Option locations were fully randomized on each trial, and trial order was counterbalanced across participants to minimize order effects.

Accuracy was scored across all 18 trials for every participant. A correct response was defined as selecting the high-ortho string alone. To examine response times and detailed eye-movement patterns, **additional inclusion criteria** was applied: only trials in which the child selected the high-ortho string were retained, and only children with at least 10 such high-ortho selections contributed to these analyses.

For response-time analyses, the latency from trial onset to the participant's mouse click on the high-ortho string was considered. Eye-tracking metrics were derived from the same high-ortho trials and included total fixation count, proportion of fixations to each stimulus type, total dwell

time, and proportion of dwell time. This report includes the desciptive statistics for accuracy, response time, and eye tracking metrics for each group.

Results

Descriptive Statistics

1. Overall Response Choices Chosen and Accuracy across Groups

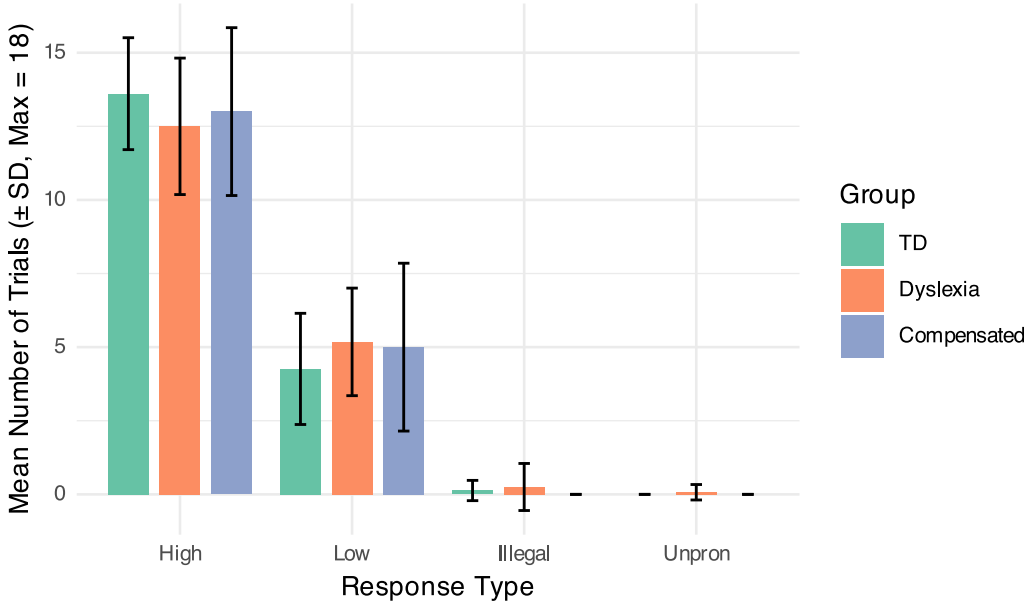
Table 1 and Figure 1 display the raw distribution of response choices across groups. All participants met the task accuracy inclusion criterion. The final sample included 23 children with typical development (TD), 28 with dyslexia, and 17 with compensated dyslexia.

Table 1: Distribution of Response Choices by Group

Mean (SD) Number of Trials per Response Type (out of 18)

Group	High Orthographic	Low Orthographic	Illegal	Unpronounceable
TD	13.61 (1.9)	4.26 (1.89)	0.13 (0.34)	0 (0)
Dyslexia	12.5 (2.32)	5.18 (1.83)	0.25 (0.8)	0.07 (0.26)
Compensated	13 (2.85)	5 (2.85)	0 (0)	0 (0)

Figure 1: Average Number of Trials Each Response Type Was Chos



Descriptive data suggest that all three groups showed a strong preference for selecting the high orthographic (high-ortho) option when asked to choose the word that looked most like a real English word. On average, children with typical development (TD) selected high-ortho items slightly more often than children with dyslexia or those with compensated dyslexia. However, the differences across groups were relatively small, and given the descriptive nature of this report, no firm conclusions can be drawn about statistical significance.

Low orthographic (low-ortho) items were selected more frequently by children in the dyslexia and compensated groups than by TD children. While all three groups demonstrated some sensitivity to legal but less typical orthographic patterns, it is possible that children with dyslexia were more likely to consider these options plausible.

Illegal and unpronounceable non-words were selected infrequently across all groups. This pattern suggests that, regardless of reading status, most children were able to avoid selecting items that clearly violated orthographic or phonological rules. Of note, no participants in the compensated group selected illegal or unpronounceable items, although this finding should be interpreted cautiously given the small number of trials per participant and the descriptive nature of the analysis.

Overall, these response patterns point to generally preserved sensitivity to orthographic structure across groups, with some variation in how children weighed high- versus low-probability forms.

Participants included for the response time and eye tracking analyses

Based on the inclusion criteria, a final sample of 63 participants (TD: n = 23, Dyslexia: n = 26, Compensated: n = 14) were included.

Only trials where “High_Ortho” was selected were part of the analyses. Each participant contributed between 10 and 17 trials, resulting in a total of 852 trials (TD = 313, Dyslexia = 335, Compensated = 204). Table 2 presents the five-number summary of included trials per participant by group.

Table 2: High-Ortho Trials per Participant

Five-number summary of included trials per participant by group

Group	Min	Q1	Median	Q3	Max	Mean
TD	10.00	12.00	14.00	15.00	17.00	13.61
Dyslexia	10.00	11.00	13.00	14.00	17.00	12.88
Compensated	10.00	11.50	14.00	15.50	17.00	13.60

---

2. Response Time (RT) across Groups

Response times were analyzed for trials in which participants selected the "High" probability orthographic non-word. Descriptive statistics by group are presented in Table 3.

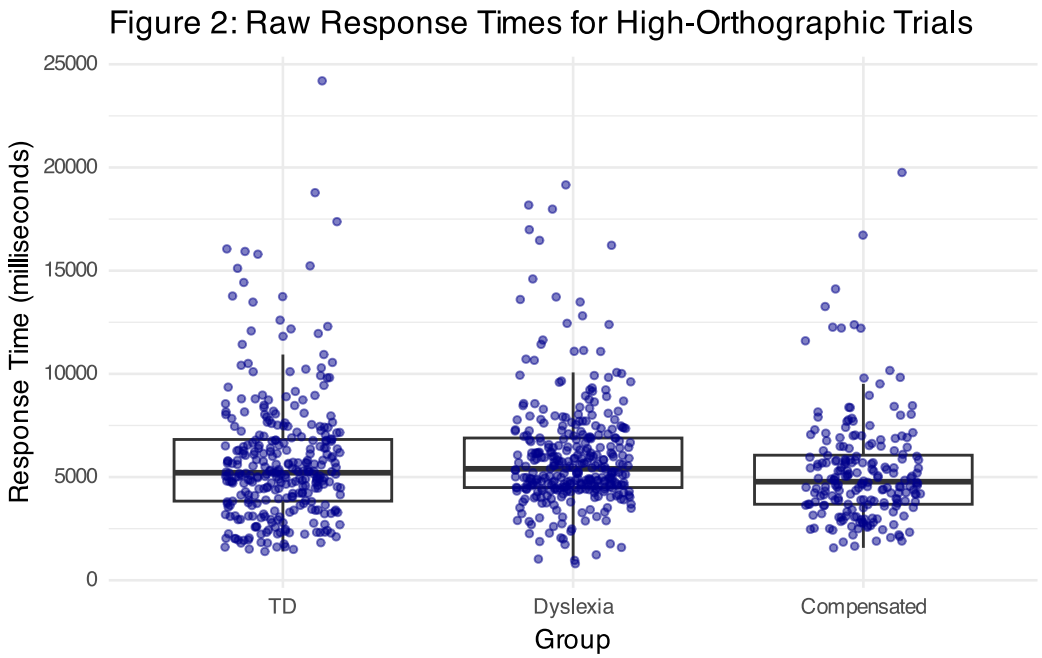
Table 3: Participant-Level Response Time Summary

Each participant's mean response time (in milliseconds)

Group	n	Mean	SD	Min	Q1	Median	Q3	Max
TD	23	5799.01	2063.3	2881.71	4196.05	5769.18	7064.80	9874.92
Dyslexia	26	6007.25	1563.4	2282.21	5368.37	5787.91	6861.43	9402.08
Compensated	15	5341.85	1694.3	3039.88	4314.46	4974.70	6121.69	8452.75

On average, participants across all groups took between 5–6 seconds to select high-orthographic items. The dyslexia group showed the longest mean response time (M = 6007 ms, SD = 1563), followed by the TD group (M = 5799 ms, SD = 2063), while the compensated group responded the fastest (M = 5341 ms, SD = 1694). Although these differences are descriptive and not statistically tested, the shorter response times observed in the compensated group may suggest greater efficiency in orthographic decision-making. This could reflect the benefit of effective intervention, as children in this group may have developed strategies that allow them to process orthographic information more quickly and consistently. Supporting this idea, the compensated group also showed a narrower interquartile range (Q1–Q3), indicating less variability in their performance compared to the other groups. Still, caution is warranted, and further analyses would be needed to determine whether these differences are statistically meaningful or reflect broader group-level patterns.

A visual inspection of the raw response time data (Figure 2) revealed a small number of unusually long trials (e.g., exceeding 20 seconds), which could reflect momentary task disengagement or other non-task-related factors.



As shown in Figure 2, response time data included a number of high-magnitude values across all groups, with some responses exceeding 20,000 milliseconds. These extreme values may reflect moments of inattention, distraction, or hesitation, and should be interpreted with caution. Prior to conducting inferential statistical analyses, it will be important to examine the distribution of response times more closely and consider appropriate handling of outliers—for example, through winsorization, trimming, or mixed-effects modeling that accounts for trial-level variability.

---

3. Eye Tracking Metrics across Groups - Fixation Count

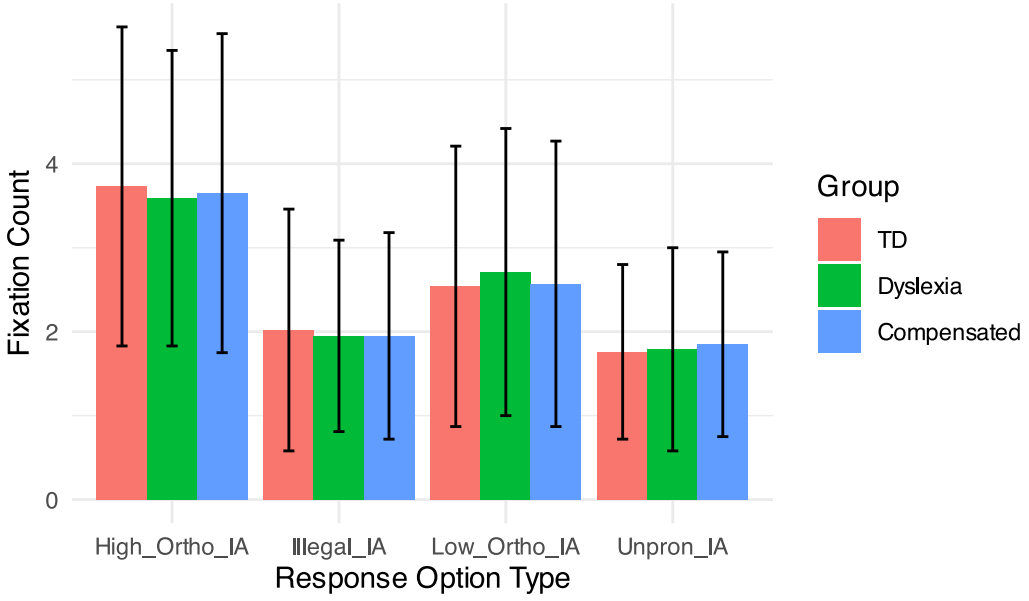
Table 5 and Figure 3 enumerate the mean (SD) number of fixations per response option type by group.

Table 5: Number of Fixations on Each Option Type

Mean (SD) Fixation Count by Group and Item Type

Group	High Ortho	Illegal Ortho	Low Ortho	Unpronounceable Ortho
TD	3.73 (1.9)	2.02 (1.44)	2.54 (1.67)	1.76 (1.04)
Dyslexia	3.59 (1.76)	1.95 (1.14)	2.71 (1.71)	1.79 (1.21)
Compensated	3.65 (1.9)	1.95 (1.23)	2.57 (1.7)	1.85 (1.1)

Figure 3: Mean Number of Fixations by Response Option Type

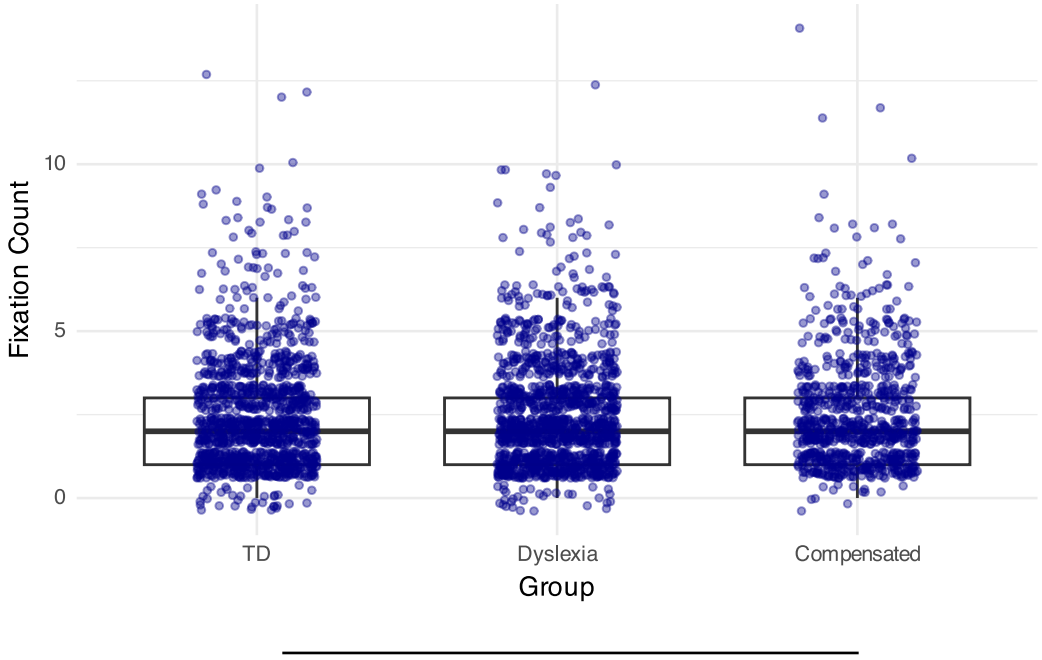


Across all groups, the high-orthographic (high-ortho) items received the greatest number of fixations, with mean values ranging from 3.59 (Dyslexia) to 3.73 (TD). Low-orthographic (low-ortho) items received the second-highest number of fixations, followed by illegal and unpronounceable options. The relatively lower fixation counts on illegal and unpronounceable items suggest that participants—regardless of group—spent less time visually inspecting these clearly implausible forms.

Fixation patterns were broadly similar across the three groups, with only minor differences in means. For example, children with dyslexia showed slightly more fixations on low-ortho items (M = 2.71) compared to TD (M = 2.54) and compensated (M = 2.57) peers. This may indicate a broader or more exploratory evaluation strategy, possibly reflecting greater uncertainty when differentiating among legal letter sequences. However, given the descriptive nature of these data and the small group differences, such interpretations remain preliminary.

Visual inspection of trial-level fixation count distributions (Figure 4) revealed the presence of outliers in all three groups. These outliers warrant further examination and appropriate handling to ensure robustness of subsequent statistical analyses.

Figure 4: Raw Fixation Counts for High-Ortho Trials



4. Eye Tracking Metrics across Groups - Proportion of Fixations

Table 6 and Figure 5 display the proportion of fixations allocated to each response option type, averaged across all included trials and participants within each group.

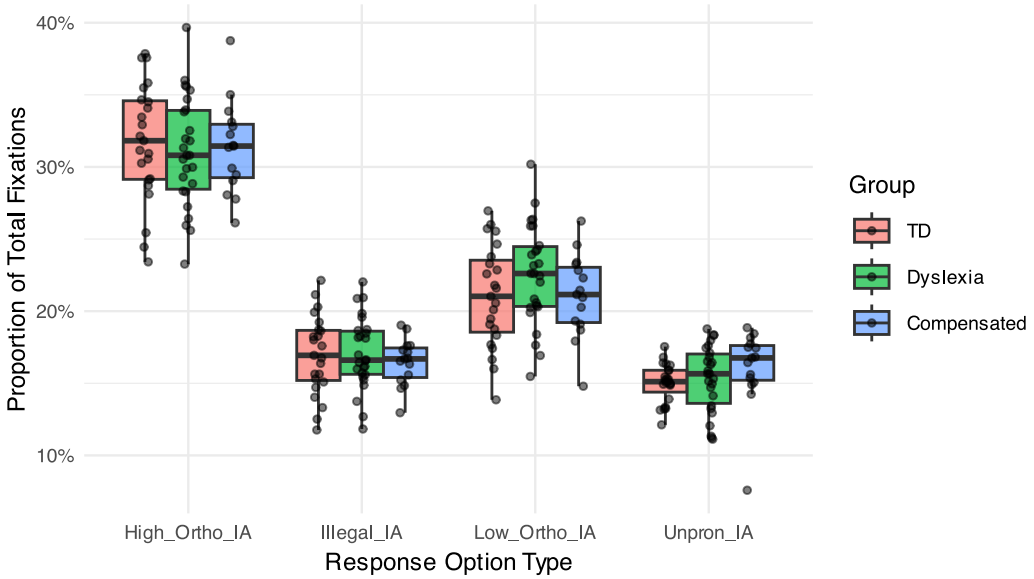
Table 6: Proportion of Fixations on Each Option Type

Mean (SD) Participant-Level % Fixation by Group and Item Type

Group	High Ortho	Illegal Ortho	Low Ortho	Unpronounceable Ortho
TD	0.32 (0.04)	0.17 (0.03)	0.21 (0.04)	0.15 (0.01)
Dyslexia	0.31 (0.04)	0.17 (0.02)	0.23 (0.03)	0.15 (0.02)
Compensated	0.31 (0.03)	0.16 (0.02)	0.21 (0.03)	0.16 (0.03)

Figure 5: Participant-Level % Fixations by Response Option Type

Each point represents a participant's average across High-Ortho trials



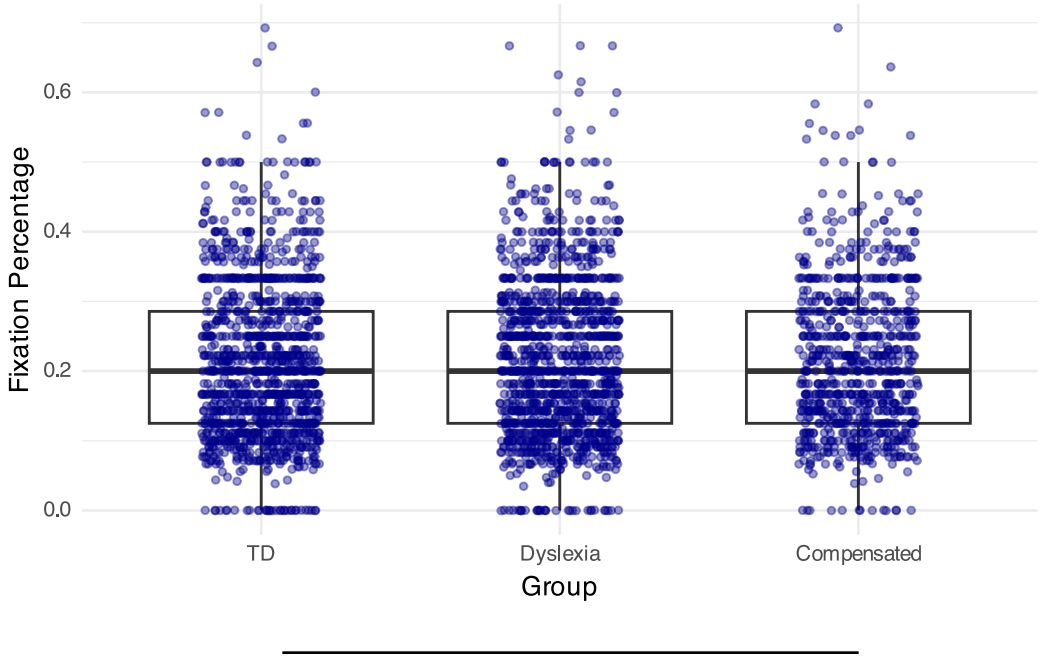
Across all three groups, participants allocated the largest proportion of their fixations to the high-orthographic (high-ortho) items during trials where they selected the high-ortho response. Mean fixation proportions for high-ortho items were nearly identical across groups, ranging from 0.31 to 0.32, with relatively low variability. This suggests that these items consistently drew visual attention, regardless of reading profile.

Participants also allocated fixations to low-orthographic and illegal items, though to a lesser extent. On average, children with dyslexia showed slightly higher fixation proportions on low-ortho items (M = 0.23) compared to the TD and compensated groups (M = 0.21), potentially reflecting broader processing of plausible letter strings or increased uncertainty during decision-making. However, these group differences were small and fall within a narrow range, and thus should be interpreted with caution.

Fixation proportions for illegal and unpronounceable options were lowest across all groups, with means between 0.15 and 0.17. This suggests that participants tended to allocate minimal attention to clearly implausible or unpronounceable forms.

As shown in Figure 6, trial-level fixation percentages on high-orthographic items varied widely within and across groups, with several outlier trials showing exceptionally high or low visual attention. These outliers should be examined prior to further analysis to ensure the reliability of gaze-based measures.

Figure 6: Raw Fixation Percentages for High-Ortho Trials



5. Eye Tracking Metrics across Groups - Dwell Time

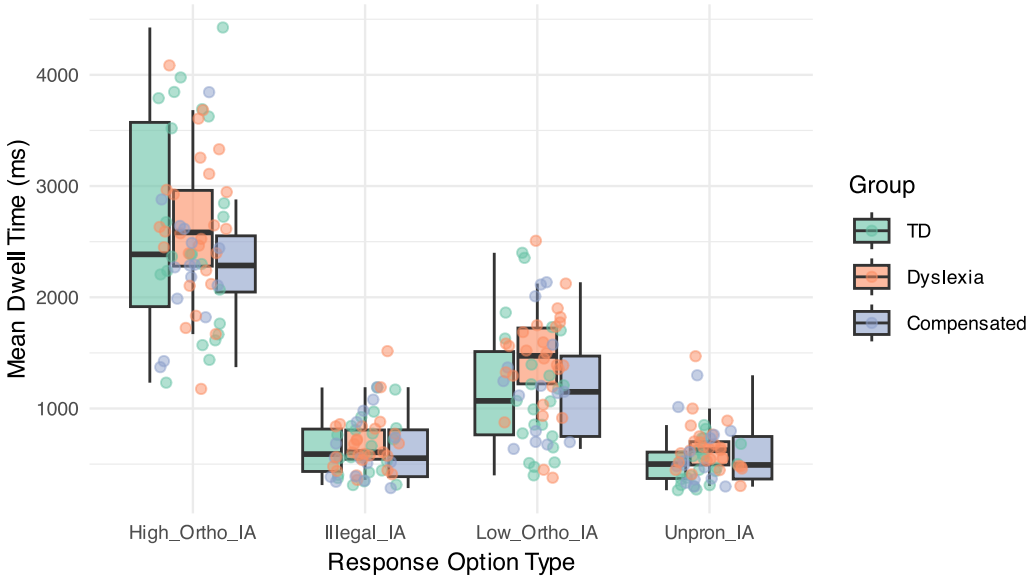
Table 7: Dwell Time on Each Option Type

Participant-level Mean (SD) Dwell Time in ms by Group and Item Type

Group	High Ortho	Illegal Ortho	Low Ortho	Unpro-nounceable Ortho
TD	2624.04 (929.03)	654.96 (257.4)	1168.88 (566.98)	508.58 (174.57)
Dyslexia	2617.92 (661.87)	680.46 (253.63)	1424.66 (473.94)	644.69 (230.76)
Compen-sated	2310.37 (597.45)	629.37 (286.96)	1237.77 (519.93)	589.8 (287)

Figure 7: Participant-Level Mean Dwell Time by Response Option Type

Each point represents a participant's average across High-Ortho trials



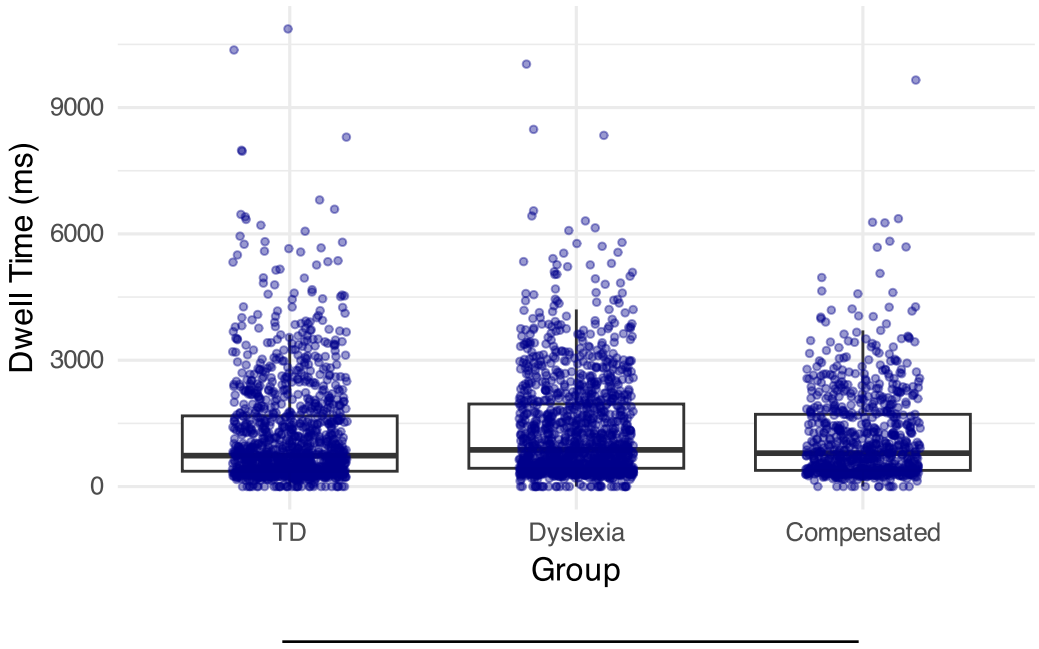
Participants in all three groups spent the most time fixating on high-orthographic (high-ortho) items, with mean dwell times ranging from approximately 2310 ms (Compensated) to 2624 ms (TD). Although the compensated group showed slightly shorter dwell times on high-ortho items than the other two groups, variability was considerable, and all groups appeared similarly engaged with these items.

Dwell time on low-orthographic (low-ortho) items was generally higher than that on illegal or unpronounceable items. The dyslexia group showed the longest dwell times on low-ortho items (M = 1424.66 ms), compared to TD (M = 1168.88 ms) and compensated participants (M = 1237.77 ms). This may suggest that children with dyslexia spent more time evaluating less typical but still legal letter sequences, possibly reflecting increased effort or broader consideration of plausible alternatives.

In contrast, illegal and unpronounceable items consistently received the shortest dwell times across all groups. Mean dwell times for illegal items ranged from 629 to 680 ms, while unpronounceable items averaged around 509–645 ms. These results indicate that participants were generally able to identify and quickly discount these less plausible options.

As shown in Figure 8, trial-level dwell time distributions on high-orthographic items revealed a number of extreme values across all groups. While the majority of trials showed relatively short dwell times, several instances of prolonged fixation suggest the presence of outliers that should be further evaluated before proceeding with inferential analyses.

Figure 8: Raw Dwell Time for High-Ortho Trials



6. Eye Tracking Metrics across Groups - First Run Dwell Time

Table 9 displays the average first-run dwell time (in seconds) for each response option type, based on participant-level means. First-run dwell time reflects the amount of time participants spent looking at an item during their first pass before moving their gaze elsewhere.

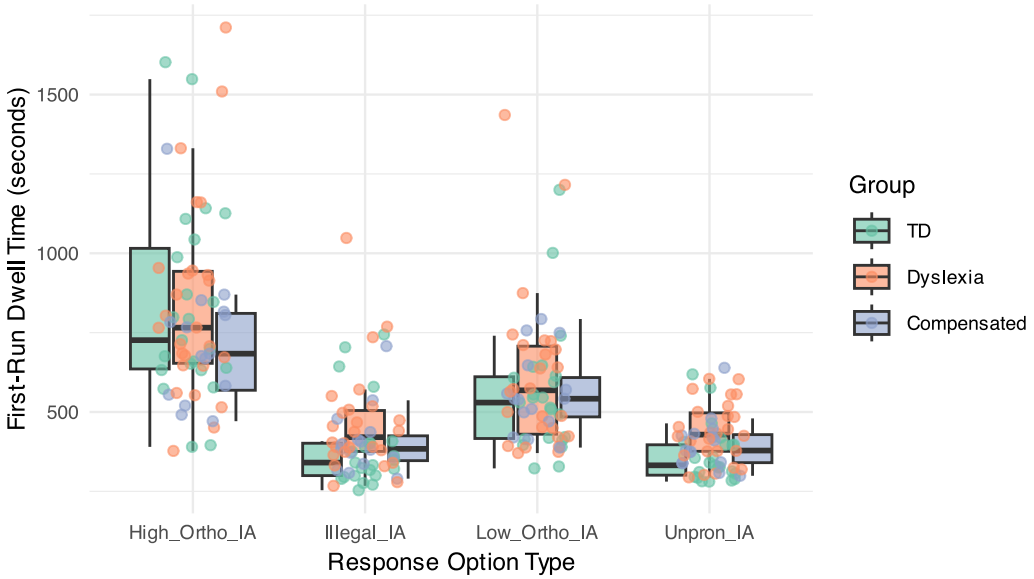
Table 8: Participant-Level First-Run Dwell Time

Each participant's mean first-run dwell time per IA (in milliseconds, raw)

Group	High Ortho	Illegal Ortho	Low Ortho	Unpro-nounceable Ortho
TD	831.21 (315.44)	390.28 (139.61)	558.15 (204.14)	363.67 (89.61)
Dyslexia	844.75 (318.32)	465.79 (168.2)	612.81 (253.98)	436.89 (93.53)
Compen-sated	724.71 (213.78)	408.1 (105.17)	559.79 (126.15)	400.45 (84.21)

Figure 9: Participant-Level First-Run Dwell Time by Response Option Type

Each point represents a participant's mean across High-Ortho trials



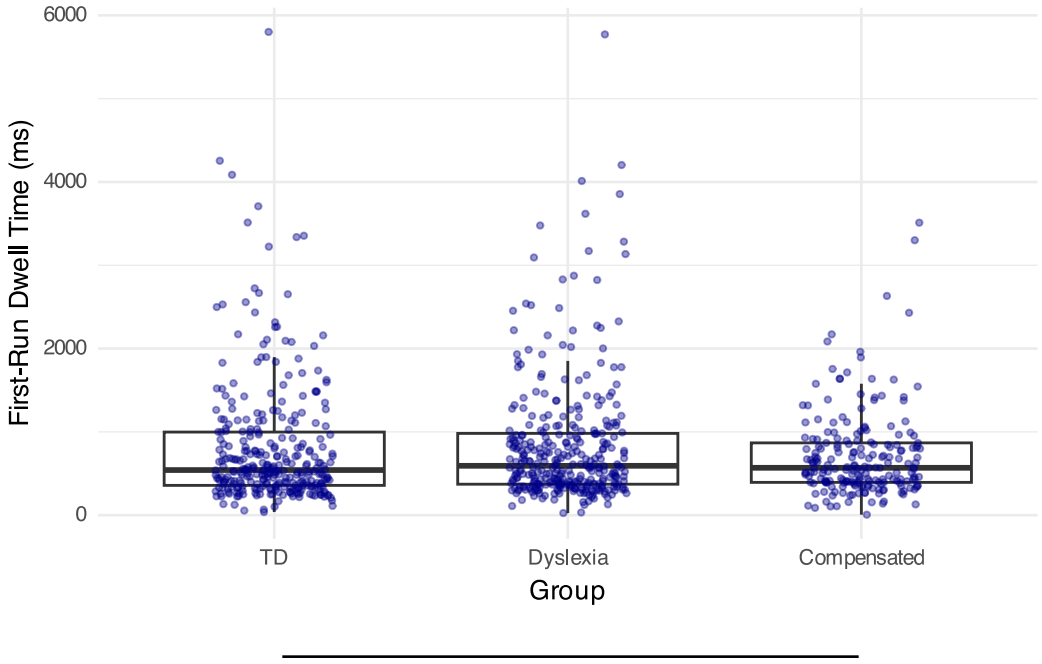
As shown in Table 8 and Figure 9, participants across all groups spent the most time in their initial pass fixating on the high-orthographic (high-ortho) items, with mean first-run dwell times ranging from approximately 725 to 845 milliseconds. The Dyslexia and TD groups showed slightly longer dwell times on these items compared to the Compensated group, though the differences were minimal and accompanied by considerable within-group variability.

For low-orthographic and illegal items, first-run dwell times were generally shorter. Notably, children with dyslexia showed slightly longer dwell times on both low-ortho (M = 613 ms) and illegal (M = 466 ms) items, potentially reflecting a more extended initial evaluation process for forms that were plausible but less typical or partially rule-violating. In contrast, the Compensated and TD groups showed slightly lower means and greater consistency across participants.

Unpronounceable items elicited the shortest first-run dwell times across all groups (M = 364–437 ms), suggesting that these clearly implausible stimuli were quickly disregarded during early visual processing. This pattern is consistent with expectations, given the lack of both phonological and orthographic plausibility in these forms.

When visually exploring the trial-level data (Figure 10), a number of outliers were observed across participants and groups. Specifically, some trials showed unusually long first-run dwell times, exceeding 3000 milliseconds, which may reflect lapses in attention or non-task-related fixations. While most values clustered within a reasonable range, the presence of these extreme cases highlights the importance of considering data cleaning procedures (e.g., winsorization or trial-level exclusion) prior to conducting inferential analyses.

Figure 10: Trial-Level First-Run Dwell Time by Group



Conclusions and Next Steps

This report provides a descriptive overview of orthographic decision-making and visual processing patterns across children with typical development, dyslexia, and compensated dyslexia. Overall, participants across all groups demonstrated sensitivity to orthographic structure, with a clear preference for selecting high-probability orthographic forms and allocating greater visual attention to these items.

Group-level differences in response time and eye-tracking metrics were observed descriptively but were modest and varied in magnitude. Children with dyslexia generally showed slightly longer response times, higher fixation counts, and longer dwell times on low-probability items compared to their peers, suggesting possible differences in processing efficiency or decision strategies. Compensated children tended to show faster and more consistent performance, potentially reflecting the benefits of earlier intervention.

Visual exploration of trial-level data revealed the presence of outliers across all metrics, including response time, fixation count, fixation percentage, dwell time, and first-run dwell time. These extreme values may reflect lapses in attention, tracking anomalies, or genuine variability in task engagement. Prior to conducting inferential statistical analyses, it will be important to carefully examine these outliers and consider appropriate handling strategies such as winsorization, trimming, or mixed-effects modeling that can accommodate trial-level variability.

Moving forward, the next steps for this project include:

- Outlier Handling: Systematically identifying and addressing extreme values to ensure robust and interpretable statistical results.
- Inferential Analyses: Conducting group comparisons using appropriate statistical models (e.g., linear mixed-effects models) to formally test hypotheses about group differences in orthographic processing.
- Exploratory Analyses: Investigating additional factors such as participant characteristics (e.g., age, reading scores) that may account for individual differences in performance.
- Reporting and Interpretation: Situating findings within the broader literature on orthographic processing and reading development, and considering implications for intervention and assessment.