

3.1 Demographic Distribution of Participants

The cognitive psychology experiment enlisted a diverse range of participants. Their background characteristics have been visualized to provide an insight into the composition of our sample.

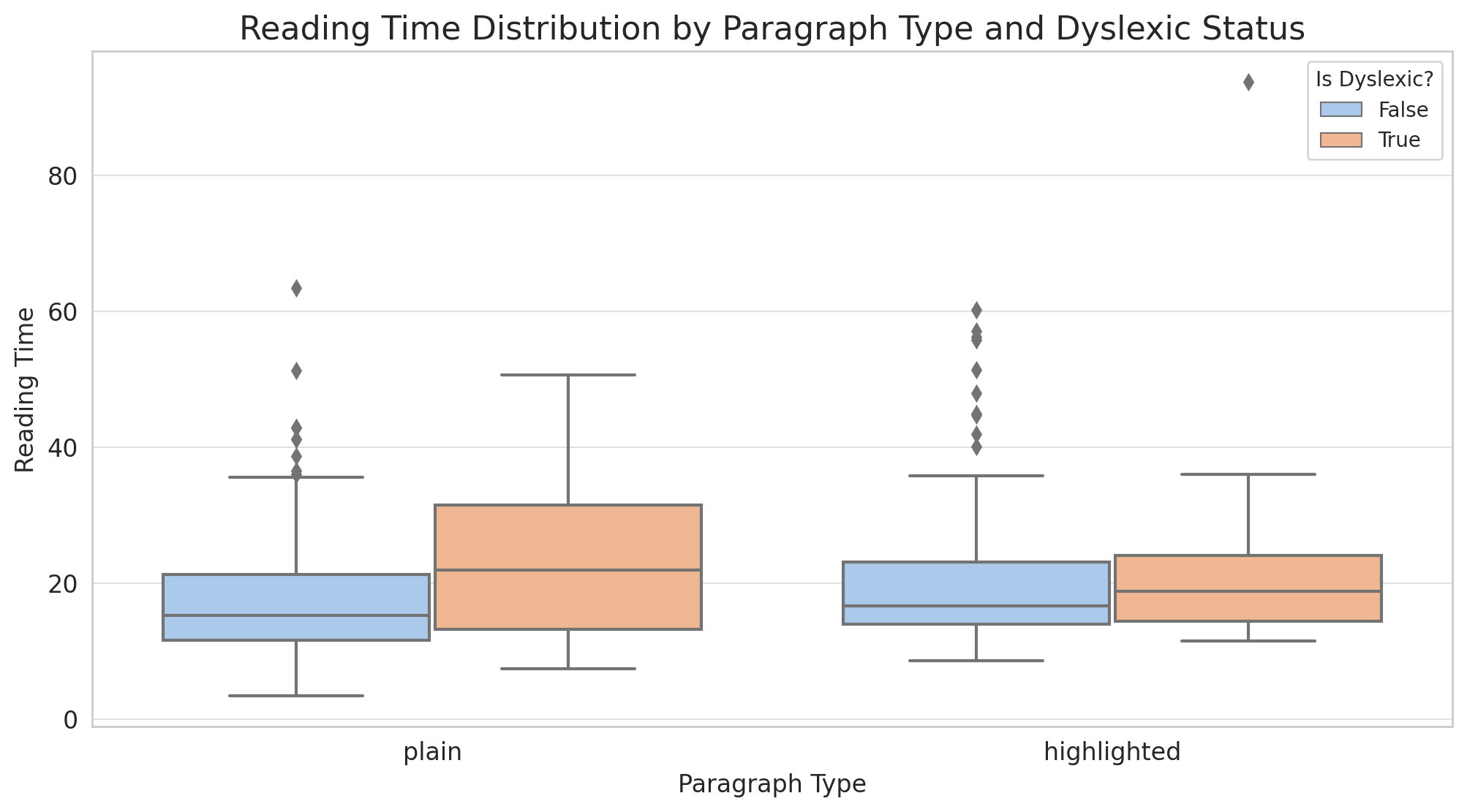
Dyslexia in Participants: The first bar chart showcases the distribution of participants based on whether they have dyslexia. A significant portion of the participants did not have dyslexia, while a smaller fraction identified as dyslexic.

Primary Language: The second bar chart illustrates that the majority of our participants do not have English as their primary language. However, a considerable number still use English as their main mode of communication.

Device Used for Participation: The third visualization reveals the preference of devices among participants. A large majority of the participants took part in the experiment using devices other than phones, suggesting a preference for larger screens or more conventional computing devices.

Age Distribution: The histogram presents the age distribution of our participants. Most participants fall within the mid-age range, with tails on the younger and older age sides. The Kernel Density Estimation (KDE) curve provides a smooth representation of this distribution, indicating the most frequent age range in the sample.

3.2 Data Visualization



The boxplot displayed below demonstrates the distribution of reading times among two distinct paragraph types: plain and highlighted. Additionally, the data is segmented based on the dyslexic status of the participants.

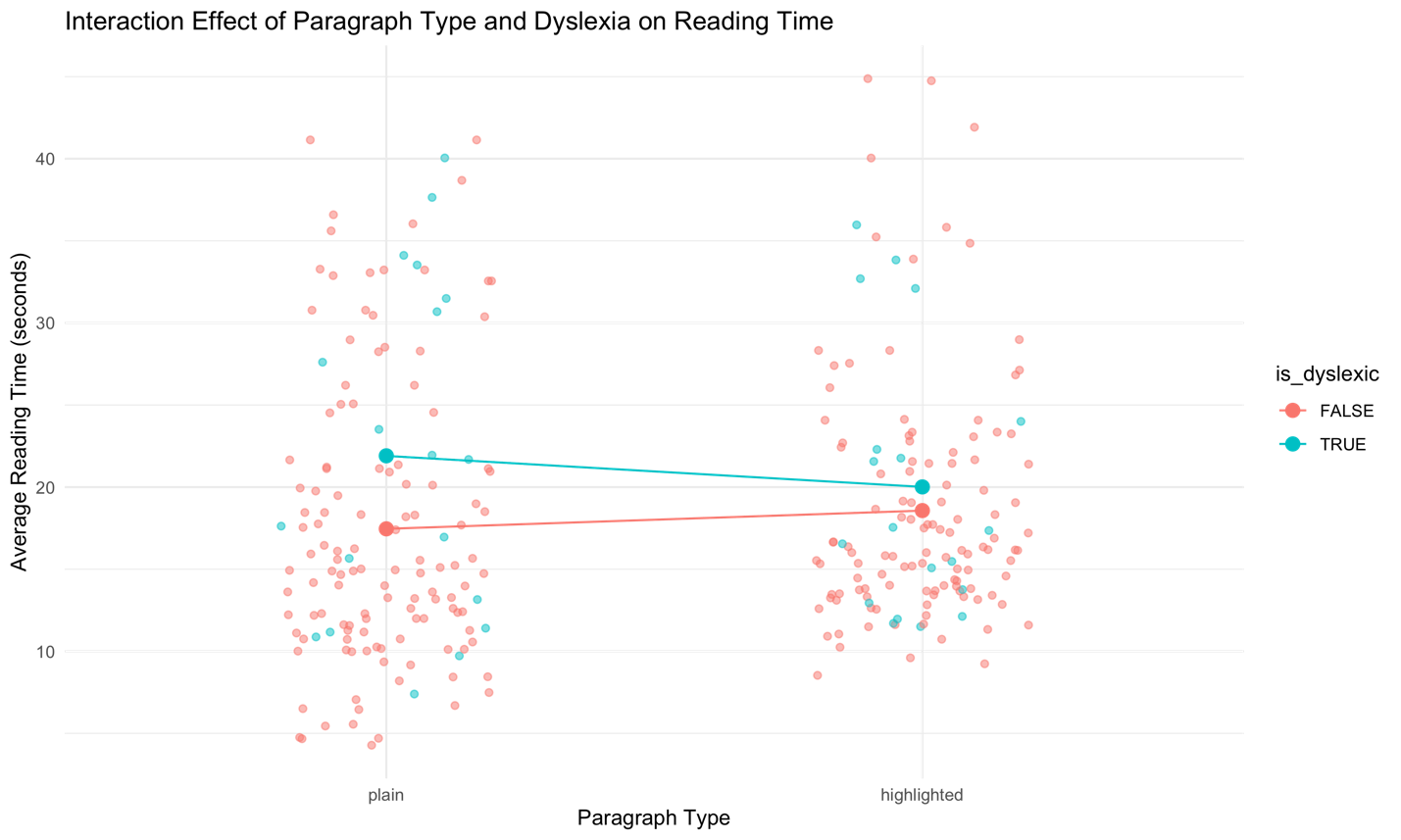
From the visual representation above, we can derive several pertinent observations:

Dyslexic vs Non-Dyslexic: Participants identified as dyslexic consistently showcased extended reading times across both paragraph types as compared to their non-dyslexic peers.

Paragraph Type Influence: The reading times for paragraphs marked as "highlighted" seem to exhibit a slight reduction when juxtaposed with the "plain" paragraph type. However, the variation remains minimal.

Presence of Outliers: A noticeable trend is the greater occurrence of outliers, especially prominent among participants with dyslexia.

Such insights offer an introductory comprehension of the data distribution, laying the groundwork for the impending mixed-effects model analysis.



3.3 LMMs

The interaction plot provided offers a comprehensive view of the combined effects of paragraph type and dyslexic status on participants' reading time.

From the interaction plot, we can discern several key observations:

Distinct Trends: The slopes of the lines for dyslexic and non-dyslexic participants differ, indicating a potential interaction between dyslexic status and paragraph type.

Dyslexic Participants: For individuals identified as dyslexic, the reading time seems to decrease when transitioning from plain to highlighted paragraphs.

Non-Dyslexic Participants: In contrast, non-dyslexic participants exhibit relatively consistent reading times regardless of the paragraph type.

Magnitude of Effect: The difference in reading times between the two paragraph types appears to be more pronounced for dyslexic participants compared to their non-dyslexic counterparts.

In essence, the interaction plot suggests that the benefit (or detriment) of highlighted paragraphs might differ based on an individual's dyslexic status. This observation emphasizes the importance of considering individual differences when designing reading materials or interventions.

4. Model Comparison and Selection

To determine the best-fitting model for our data, two linear mixed models were compared:

Model 1 (Simplified Model): This model incorporated the fixed effects of paragraph type and dyslexic status, their interaction, and a random intercept for subjects.

reading\_time∼paragraph\_type×is\_dyslexic+(1∣subject)

Model 2 (Complex Model): In addition to the fixed effects present in Model 1, this model included random slopes for both paragraph type and dyslexic status within subjects.

reading\_time∼paragraph\_type×is\_dyslexic+(1+paragraph\_type+is\_dyslexic∣subject)

Model Comparison Results:

Number of Parameters (npar): Model 2, being more complex, has more parameters (11) than Model 1 (6).

AIC and BIC: Both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are lower for Model 2, indicating a better fit to the data.

Log-Likelihood: Model 2 has a higher log-likelihood value, further suggesting it fits the data better.

Chi-Squared Test: The chi-squared test comparing the two models is significant (p < 0.001), with a value of 36.73 on 5 degrees of freedom. This suggests that the additional complexity in Model 2 provides a significantly better fit to the data.

Conclusion:

Given the aforementioned results, Model 2 (the complex model with additional random slopes) provides a significantly better fit to the data when compared to Model 1. Therefore, Model 2 is selected for further interpretation and analysis.

3.5 Detailed Analysis of the Selected Model

Based on the previous model comparison, the more complex model, which includes both fixed effects and their interaction, as well as random slopes for paragraph type and dyslexic status, was selected for further analysis.

Model Specifications:

The model is represented by the formula:

reading\_time∼paragraph\_type×is\_dyslexic+(1+paragraph\_type+is\_dyslexic∣subject)

Random Effects:

The variance for the intercept across subjects is 43.34 with a standard deviation of 6.583.

The variance for the effect of paragraph\_type (highlighted) across subjects is 43.17 with a standard deviation of 6.571. Its correlation with the intercept is -0.53.

The variance for the effect of is\_dyslexic (True) across subjects is 80.67 with a standard deviation of 8.981. Its correlation with the intercept is -0.47, and with paragraph\_type is -0.31.

The residual variance is 29.43 with a standard deviation of 5.425.

Fixed Effects:

Intercept: The estimated average reading time for non-dyslexic participants reading plain paragraphs is 17.8227 units.

Paragraph Type (Highlighted): The estimated change in reading time when switching from a plain paragraph to a highlighted paragraph for non-dyslexic participants is an increase of 0.9574 units, though this effect is not statistically significant (t = 0.620).

Dyslexic Status (True): Dyslexic participants reading plain paragraphs have an estimated increase in reading time of 5.4344 units compared to non-dyslexic participants, but this effect is also not statistically significant (t = 1.192).

Interaction Effect: The interaction between paragraph type (highlighted) and dyslexic status (True) results in an estimated decrease in reading time of 4.8811 units. This indicates that while dyslexic participants generally have longer reading times, the highlighted paragraph seems to decrease this effect, albeit the effect is not statistically significant (t = -1.196).

Conclusion:

The selected model reveals that there's an interaction between paragraph type and dyslexic status on reading time. Although the individual effects of paragraph type and dyslexic status are not statistically significant in this model, their interaction suggests a potential modulation of the effect of paragraph type by dyslexic status. This underscores the importance of considering both factors in tandem when evaluating reading times.

3.6 Analysis of Correct Response Using a Generalized Linear Mixed Model (GLMM)

To investigate the factors influencing the correct response of participants, a generalized linear mixed model (GLMM) with a binomial family (logit link) was applied. The model evaluated the effects of paragraph type, dyslexic status, and their interaction while accounting for random intercepts for subjects.

Model Specifications:

The model is given by the formula:

correct\_response∼paragraph\_type×is\_dyslexic+(1∣subject)

Model Fit:

AIC and BIC: The model's Akaike Information Criterion (AIC) is 227.9, and the Bayesian Information Criterion (BIC) is 246.6, which provide measures of the model's fit to the data relative to other potential models.

Log-Likelihood: The log-likelihood value is -108.9.

Residuals: The scaled residuals range from a minimum of -3.1272 to a maximum of 0.4597, with the majority of residuals concentrated around the median value of 0.3592.

Random Effects:

The variance of the random intercept for subjects is 0.00491, with a standard deviation of 0.07007.

Fixed Effects:

Intercept: The estimated log odds of a correct response for non-dyslexic participants reading plain paragraphs is 2.04153 (z = 7.043, p < 0.001).

Paragraph Type (Highlighted): The estimated change in log odds when switching from a plain paragraph to a highlighted paragraph for non-dyslexic participants is 0.23752, but this effect is not statistically significant (z = 0.589, p = 0.556).

Dyslexic Status (True): Dyslexic participants reading plain paragraphs have an estimated decrease in log odds of a correct response of -0.48110 compared to non-dyslexic participants, but this effect is not statistically significant (z = -0.785, p = 0.432).

Interaction Effect: The interaction between paragraph type (highlighted) and dyslexic status (True) results in an estimated change in log odds of 0.04889. This effect is not statistically significant (z = 0.053, p = 0.958).

Conclusion:

The GLMM analysis suggests that neither the paragraph type nor the dyslexic status, nor their interaction, have a statistically significant effect on the likelihood of a correct response. The only significant predictor is the intercept, indicating that other unaccounted factors might play a role in determining the correct response. Further exploration and possibly inclusion of other predictors may be warranted to better understand the factors influencing correct responses.