

The attentive online reading: do eye movement modeling examples enhance navigation and evaluation of multiple documents in English (L2)?

Contents

Preliminary Analyses	2
Outliers Detection and Replacement	2
Tests of Normality and Data Transformation	2
Main Analysis	3
Research Questions 1a and 1b	3
Figure 3	4
Figure 4	5
Research Questions 2a, 2b and 2c	5
Figure 5	6
Figure 6	7
Figure 7	8
Figure 8	9
Figure 9	10
Figure 10	11
Figure 11	12
Figure 12	13
Figure 13	14
Figure 14	15
Figure 15	16
Figure 16	17
Figure 17	18
Figure 18	19
Figure 19	20
Figure 20	21
Figure 21	22
Figure 22	23
Figure 23	24
References	24

Preliminary Analyses

Outliers Detection and Replacement

Individual fixations that lasted two standart deviations above or below each participant fixation duration mean were considered outliers and were replaced by the participant fixation duration median(see SALMERÓN; DELGADO; MASON, 2020, p. 1045). Outliers represented 4.56% of the durations.

Tests of Normality and Data Transformation

Figure 1 shows the frequency distributions of the time variables used in the study. All distributions have skewness values higher than 0.5, being right-skewed, as it is commonly observed with time variables. The variables were log-transformed to be better approximated to a normal distribution. The log-transformed variables are shown in Figure 2 with their theoretical probability distributions. Transformation resulted in improved skewness values for two of the four variables, namely, total fixation duration on serp and total fixation duration on source features. Thus, log-transformed variables were used for these two variables.

(see SALMERÓN; DELGADO; MASON, 2020, p. 1047)

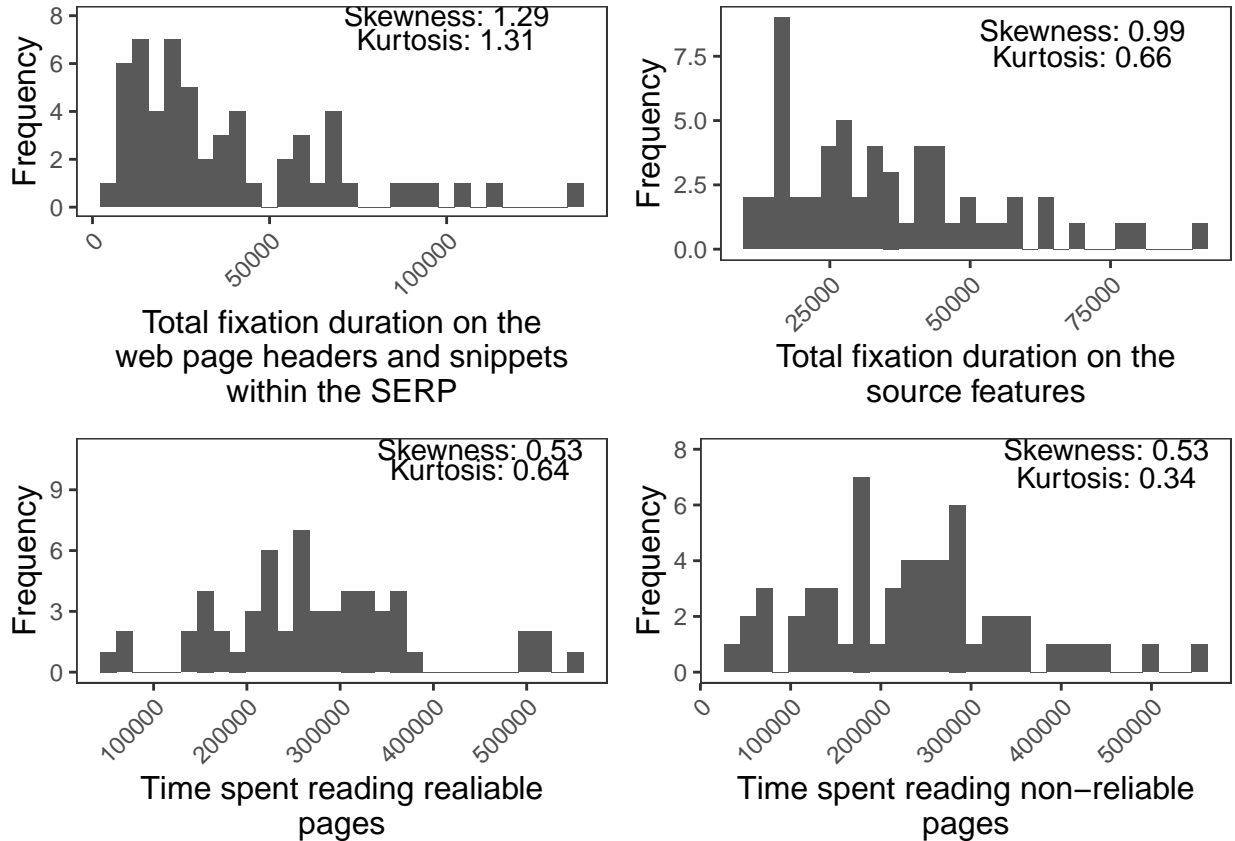


Figure 1. Histograms of time variables

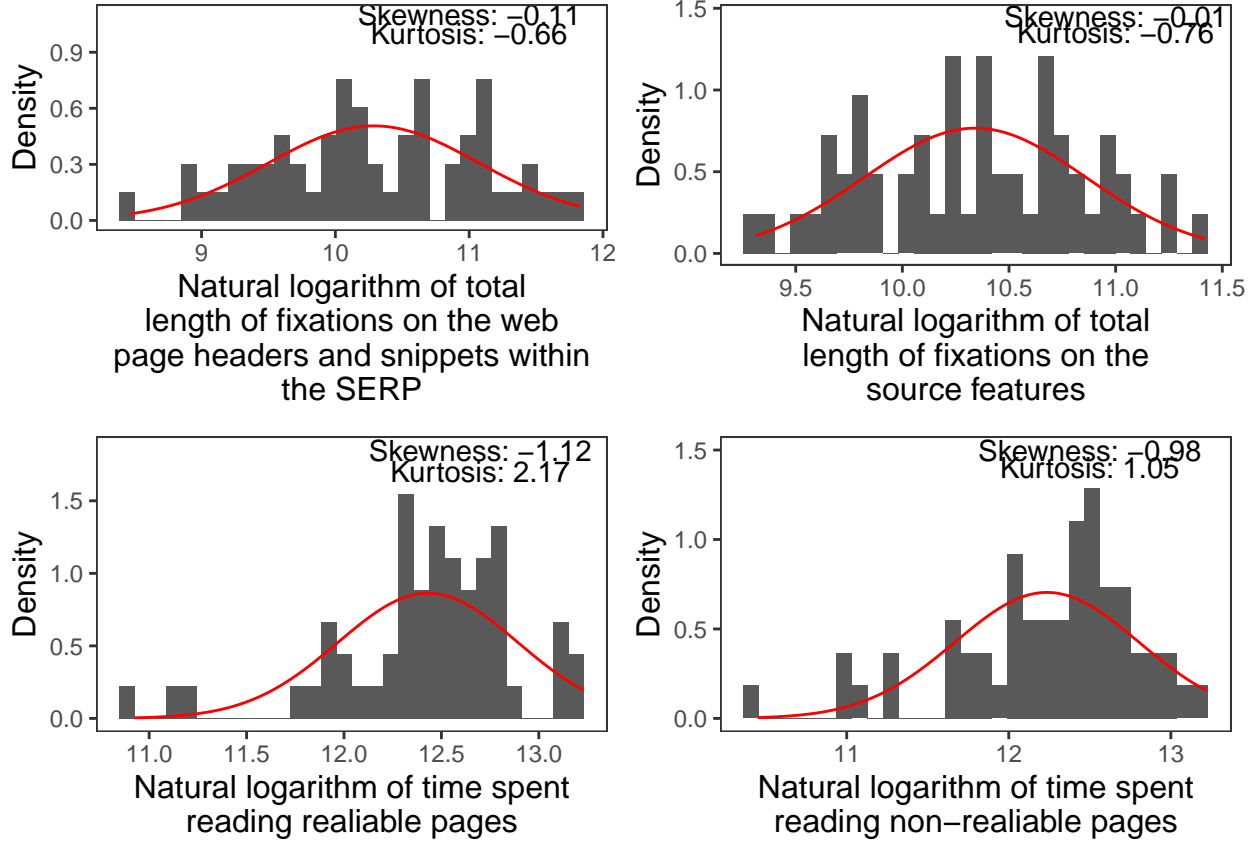


Figure 2. Histograms of transformed time variables with theoretical probability distributions

Main Analysis

Research Questions 1a and 1b

To answer the research questions RQ1a and RQ1b, two linear models were fit: the first one was a multiple linear regression model with the natural logarithm of fixation duration on SERP as response and Group (two levels: EMME and VIDEO), L2 level, and self-perceived strategic behavior as predictors; the second one was a mixed-effects model with total fixation duration on source features, on reliable and non-reliable pages, as response, and Group, L2 level, and self-perceived strategic behavior as predictors with random intercepts for participants.

The first model shows no statistically significant effect when all the possible interactions are considered. When interactions are removed from the model, it shows a statistically significant effect of Group ($\beta = -0.67, p = 0.001, 95\%CI = [-1.06, -0.28], R^2 = .19, R^2_{adjusted} = .15$) (see Table 1).

The second model shows a similar pattern: no statistically significant effects when the interactions are considered and a statistically significant effect of Group ($\beta = 56121.42, p = 0.040, 95\%CI = [2604.90, 109637.94], MarginalR^2 = .12, ConditionalR^2 = .77$) when the interactions are removed from the model (see Table 2).

Figure 3

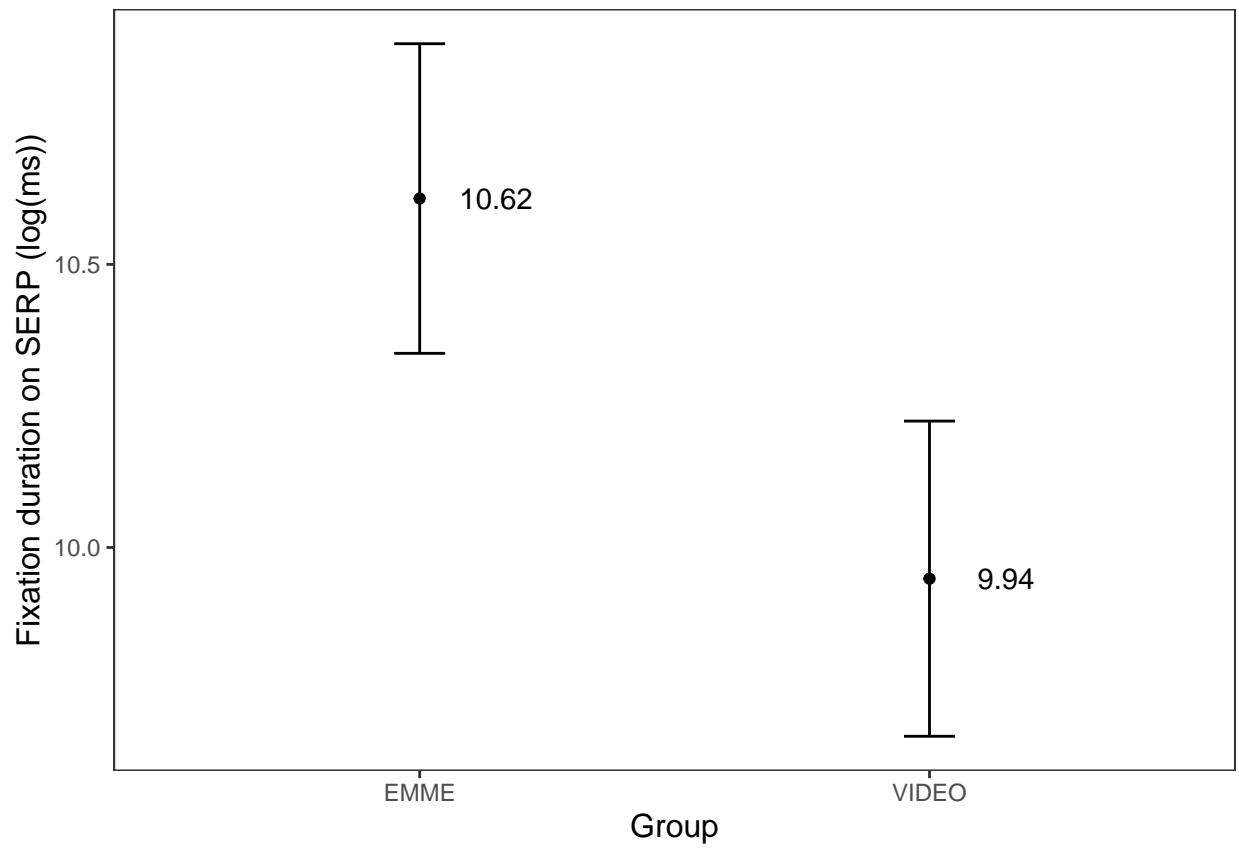


Figure 3. Log of total lenght of fixations predicted by Group

Figure 4

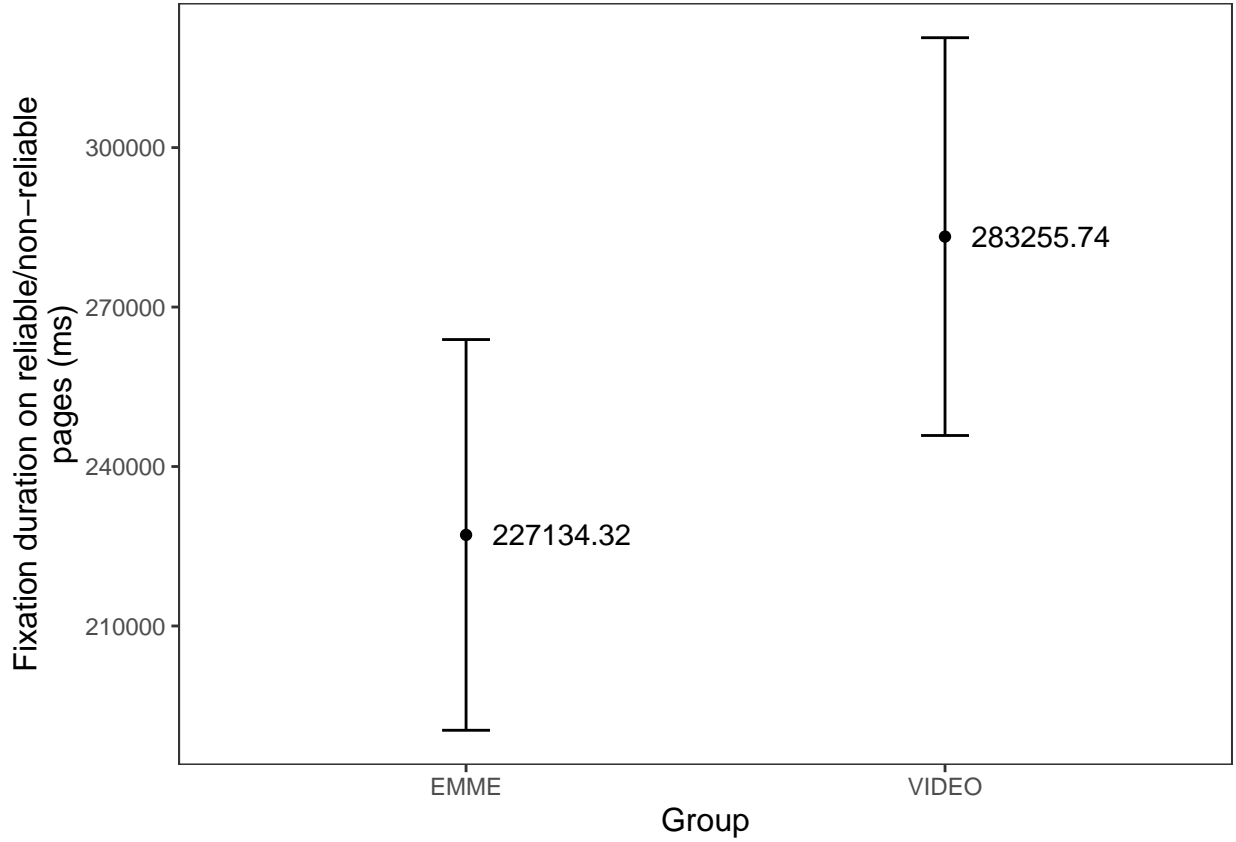


Figure 4. Fixation duration on reliable/non-reliable pages predicted by Group

Research Questions 2a, 2b and 2c

To answer the reserch questions RQ2a, RQ2b, and RQ2c, three linear models were fit: the first with the argumentation scores as response and Group, English level, Fixation duration on SERP and fixation duration on source features as predictors; the second model with smt scores as response and Group, English level, Fixation duration on SERP and fixation duration on source features as predictors; and the third model, a mixed-effects model, with scores on the pre and post tests as response, Group, English level, Fixation duration on SERP, Fixation duration on source features, and moment (pre and post) as predictors, and random intercepts for participants.

The first model ($R^2 = 0.412$, $R^2_{adjusted} = 0.196$) shows statistically significant effects of Fixation duration on SERP ($\beta = -207.00$, $p = 0.040$, $95\%CI = [-403.65, -10.36]$) and Fixation duration on source features ($\beta = -233.13$, $p = 0.025$, $95\%CI = [-435.59, -30.66]$) (Figures 5 and 6). Additionally, there are statistically significant interactions between Group and Fixation duration on source features ($\beta = -245.26$, $p = 0.044$, $95\%CI = [7.19, 483.32]$), English level and Fixation duration on source features ($\beta = 2.71$, $p = 0.047$, $95\%CI = [0.04, 5.38]$), and Fixation duration on SERP and Fixation duration on source features ($\beta = 21.17$, $p = 0.030$, $95\%CI = [2.13, 40.22]$) (Table 3 and Figures 7 to 9).

The second model ($R^2 = 0.64$, $R^2_{adusted} = 0.514$) shows statistically significant effects for all predictors and interactions (see Table 4 and Figures 10 to 23).

The third model ($MarginalR^2 = 0.167$, $ConditionalR^2 = 0.436$) shows statistically significant effects of English level ($\beta = 0.05$, $p = 0.025$, $95\%CI = [0.01, 0.10]$), Fixation duration on source features ($\beta = -0.95$, $p =$

0.020, 95%CI = $[-1.75, -0.15]$), and Moment ($\beta = -0.90, p = 0.002, 95\%CI = [-1.47, -0.33]$), see Table 5.

Figure 5

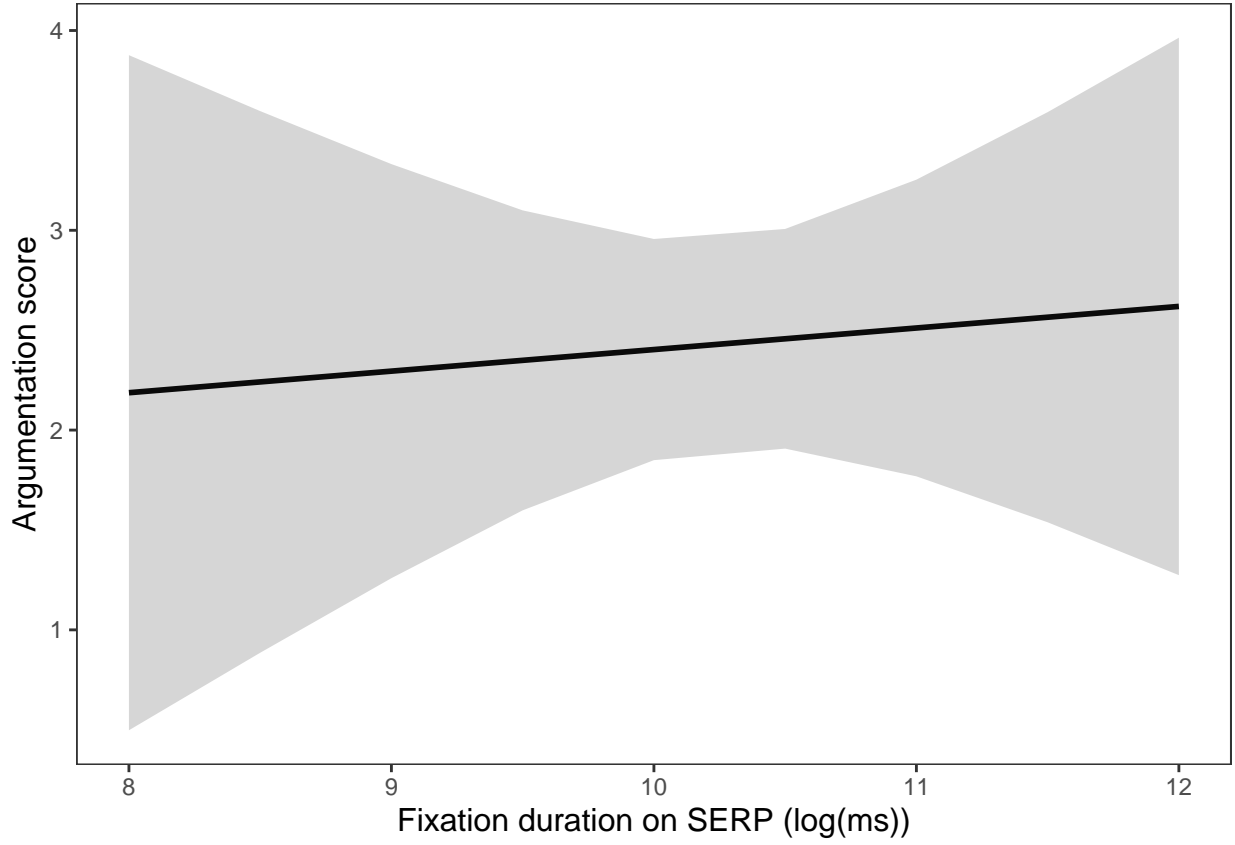


Figure 5. Argumentation scores predicted by Fixation duration on SERP

Figure 6

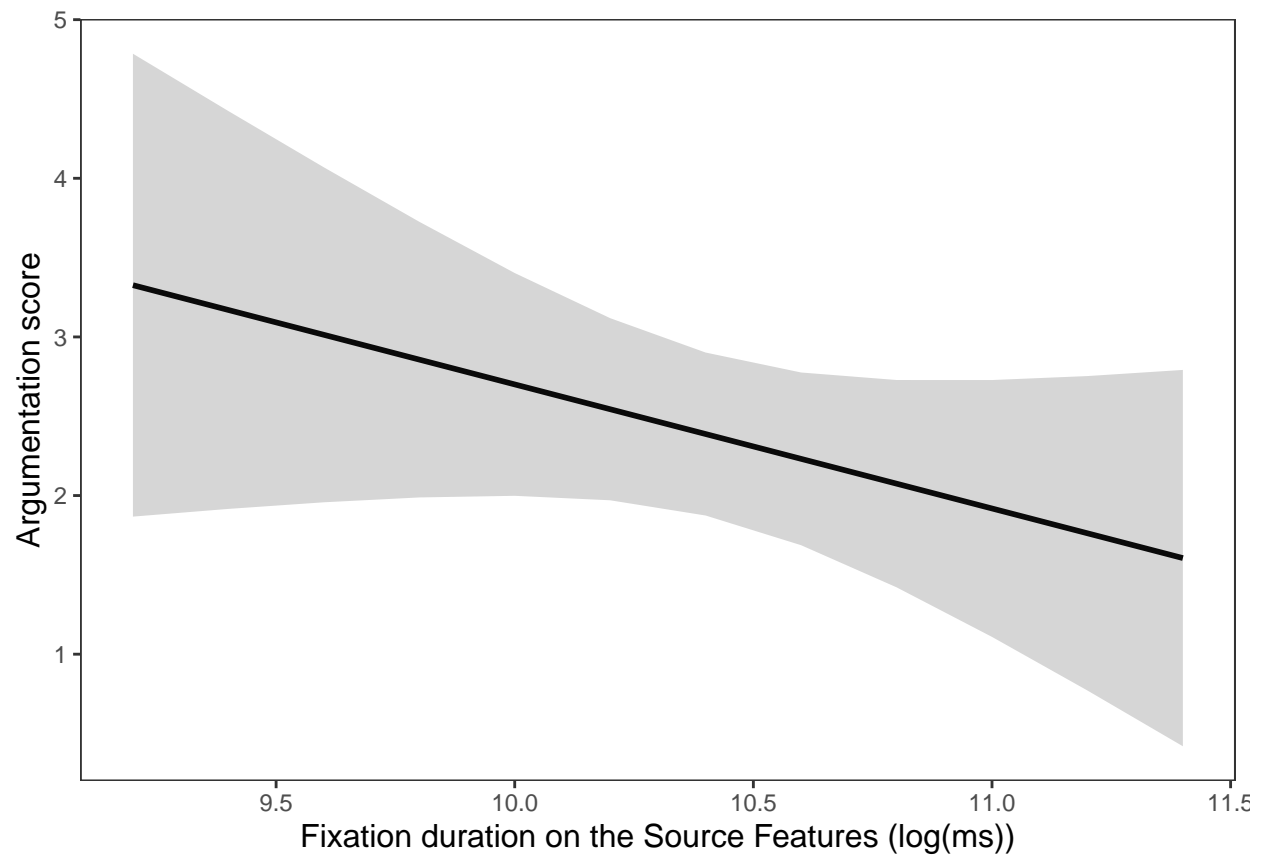


Figure 6. Argumentation score predicted by Fixation duration on the Source Features

Figure 7

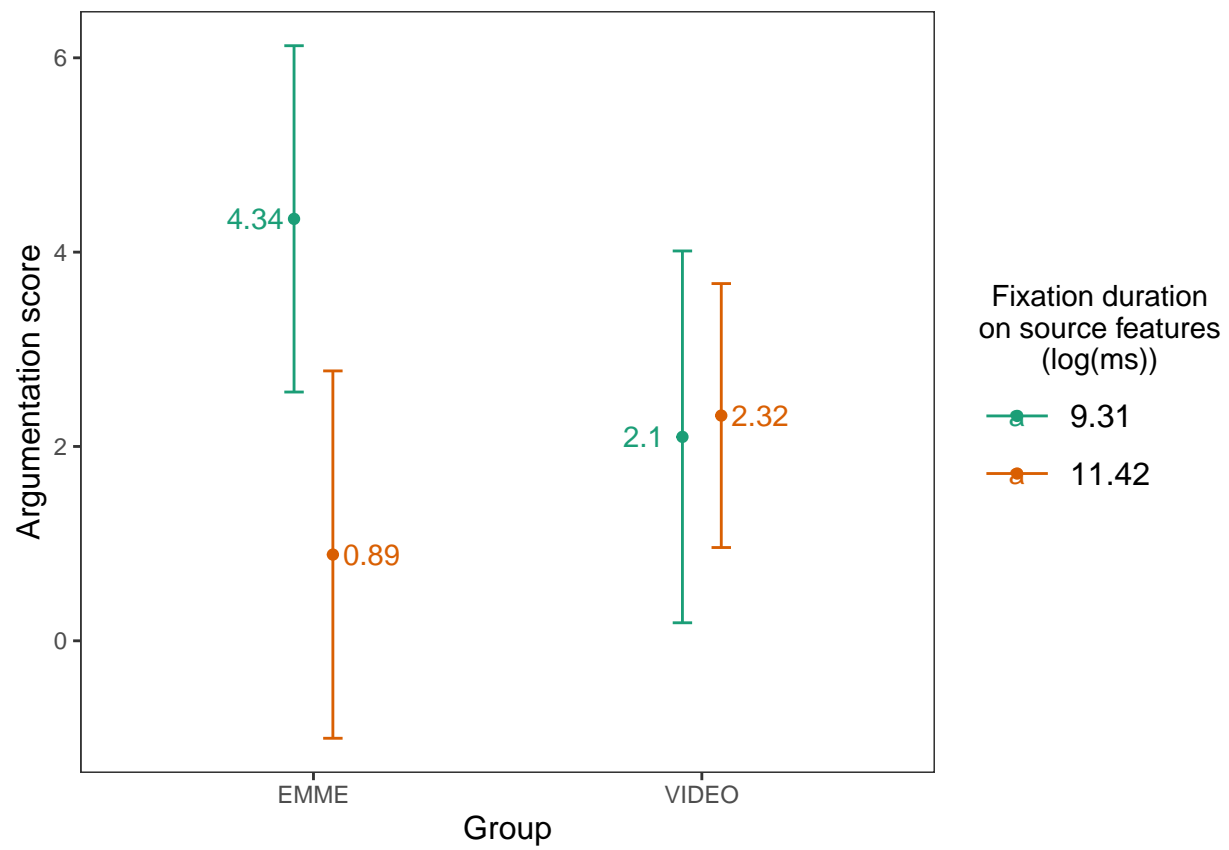


Figure 7. Interaction between Group and Fixation duration on source features

Figure 8

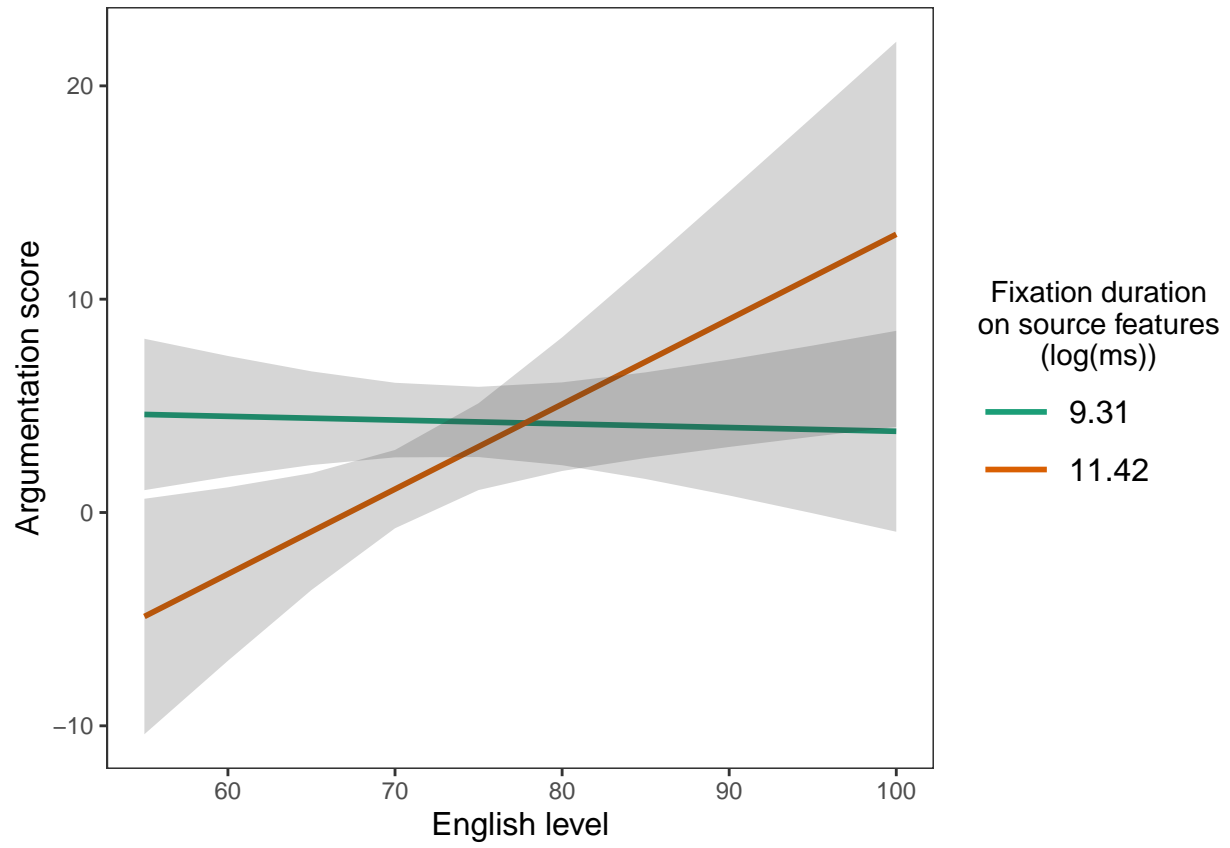


Figure 8. Interaction between English level and Fixation duration on source features

Figure 9

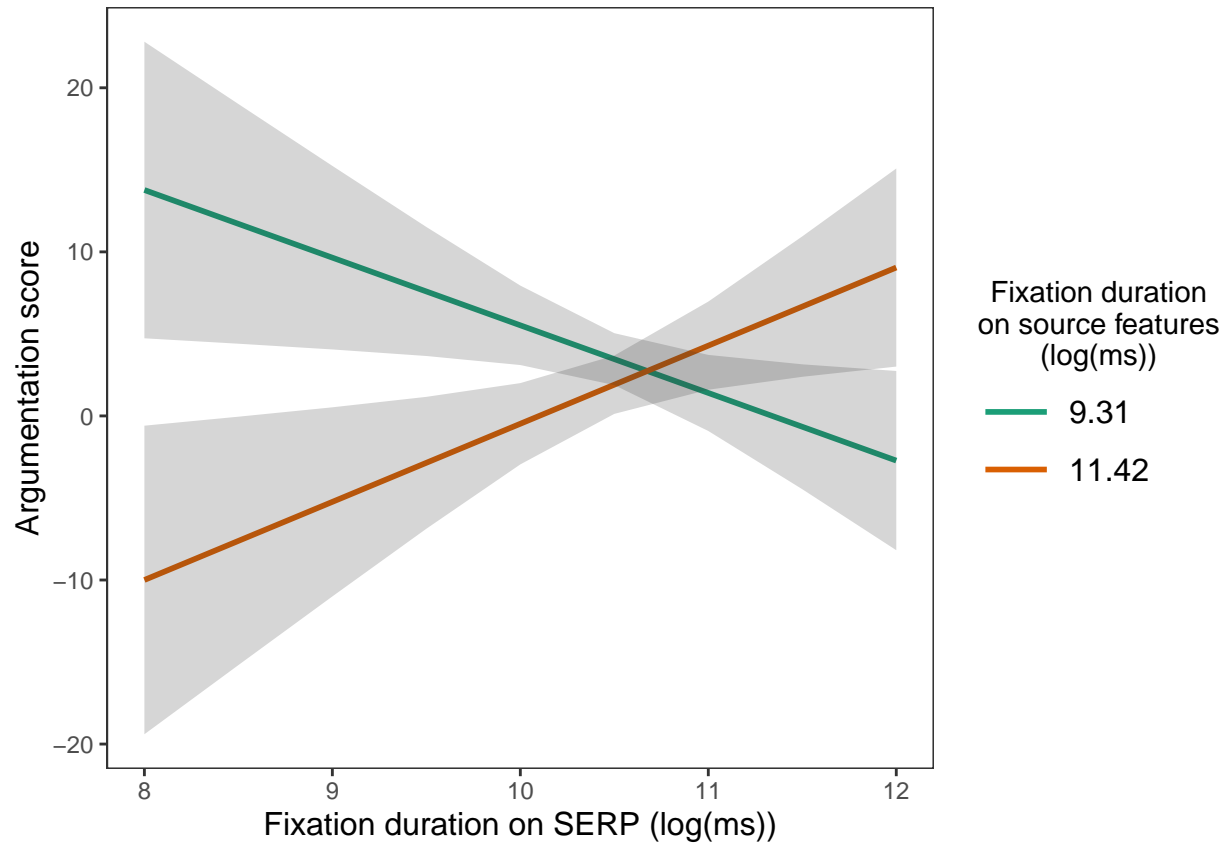


Figure 9. Interaction between Fixation duration on SERP and Fixation duration on source features

Figure 10

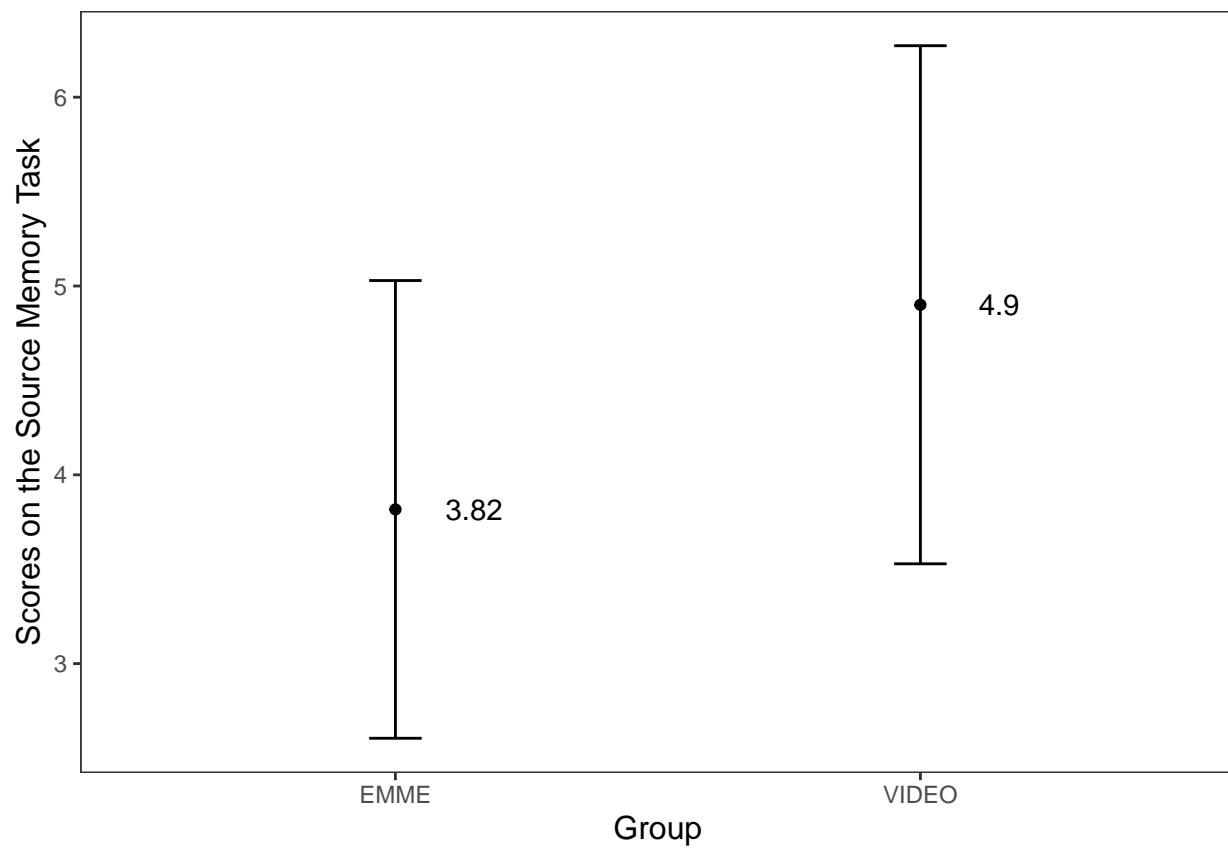


Figure 10. Scores on the Source Memory Task predicted by Group

Figure 11

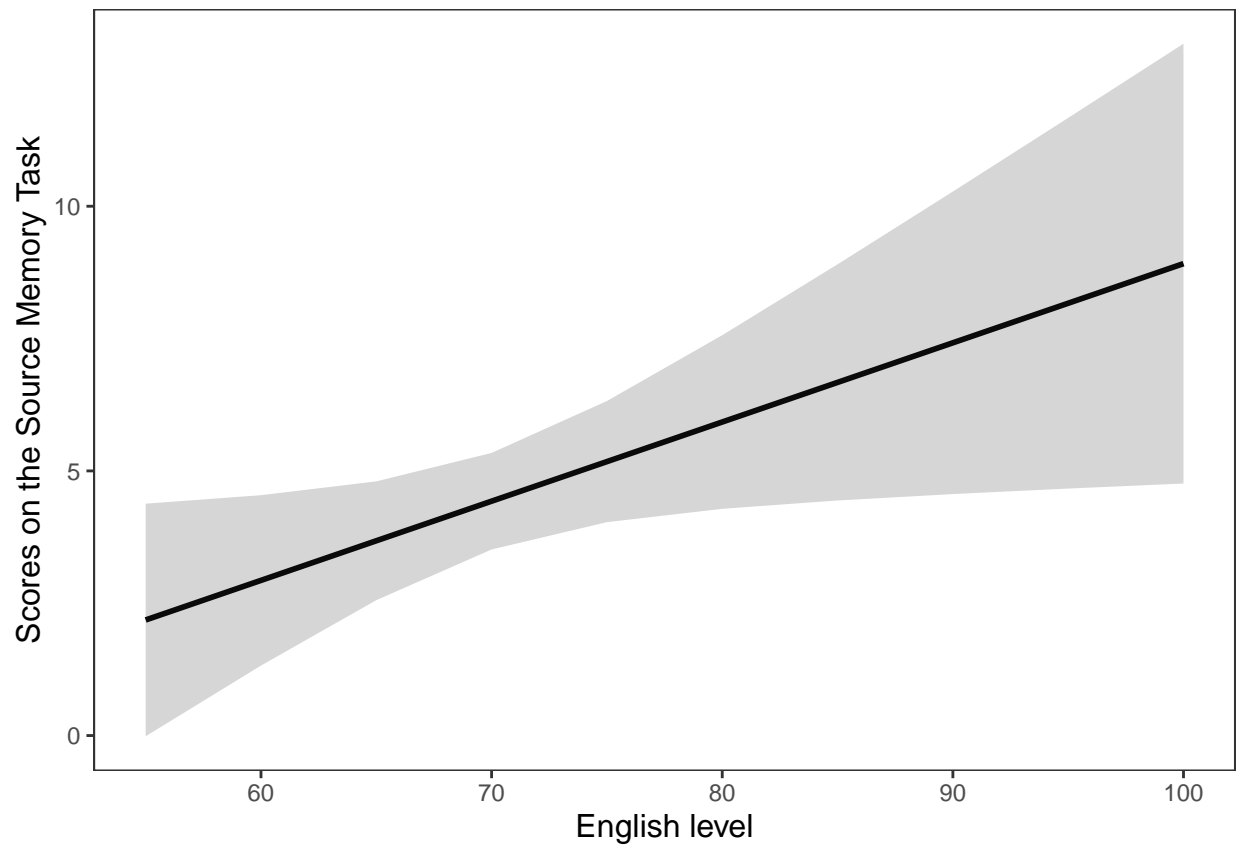


Figure 11. Scores on the Source Memory Task predicted by English level

Figure 12

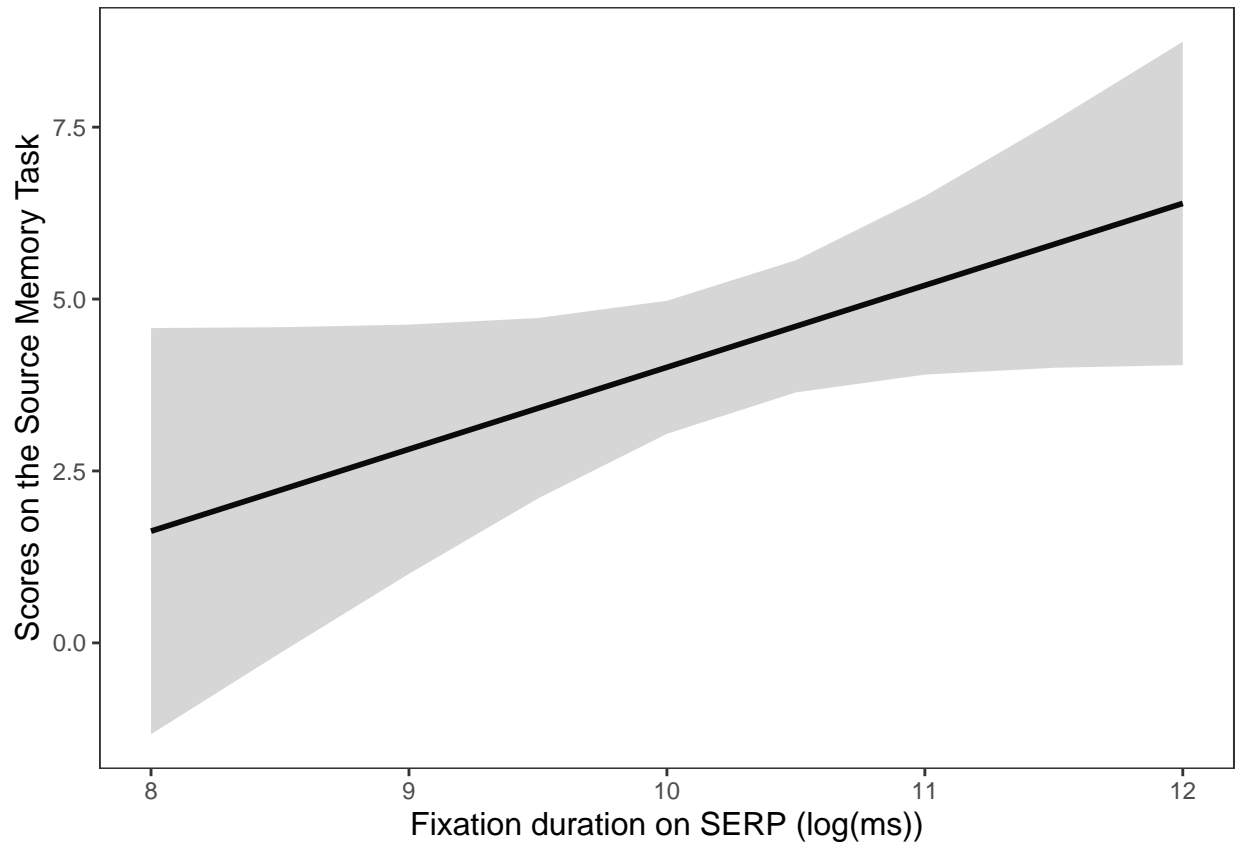


Figure 12. Scores on the Source Memory Task predicted by Fixation duration on SERP

Figure 13

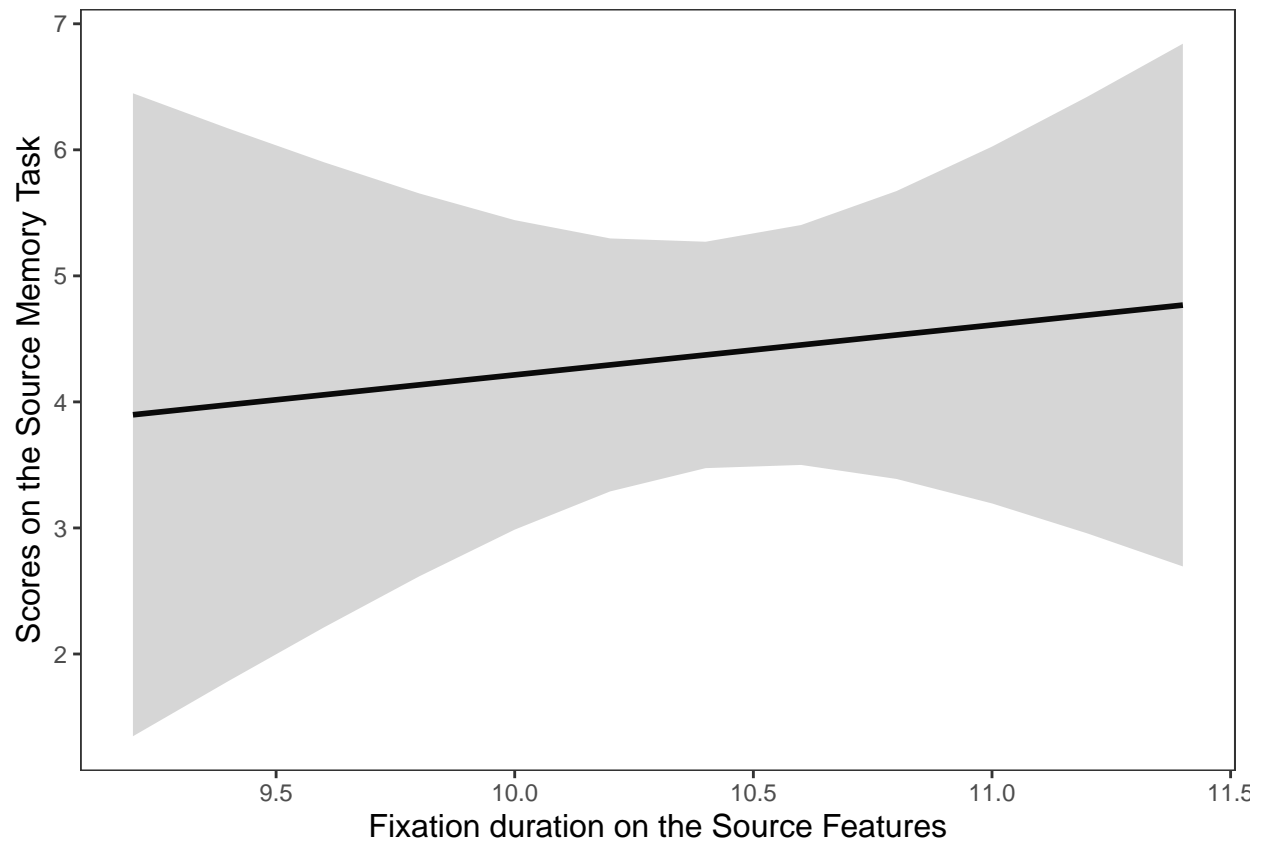


Figure 13. Scores on the Source Memory Task predicted by Fixation duration on the Source Features

Figure 14

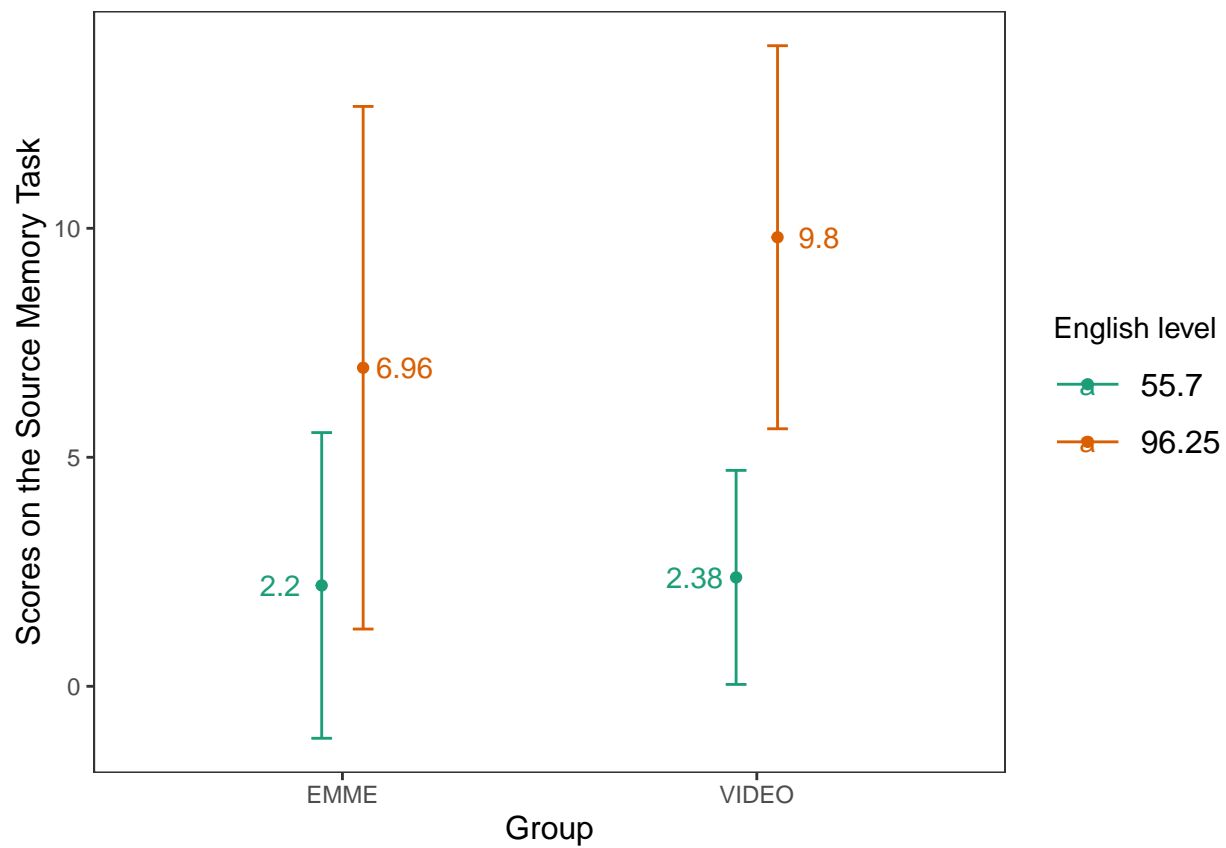


Figure 14. Interaction between Group and English level

Figure 15

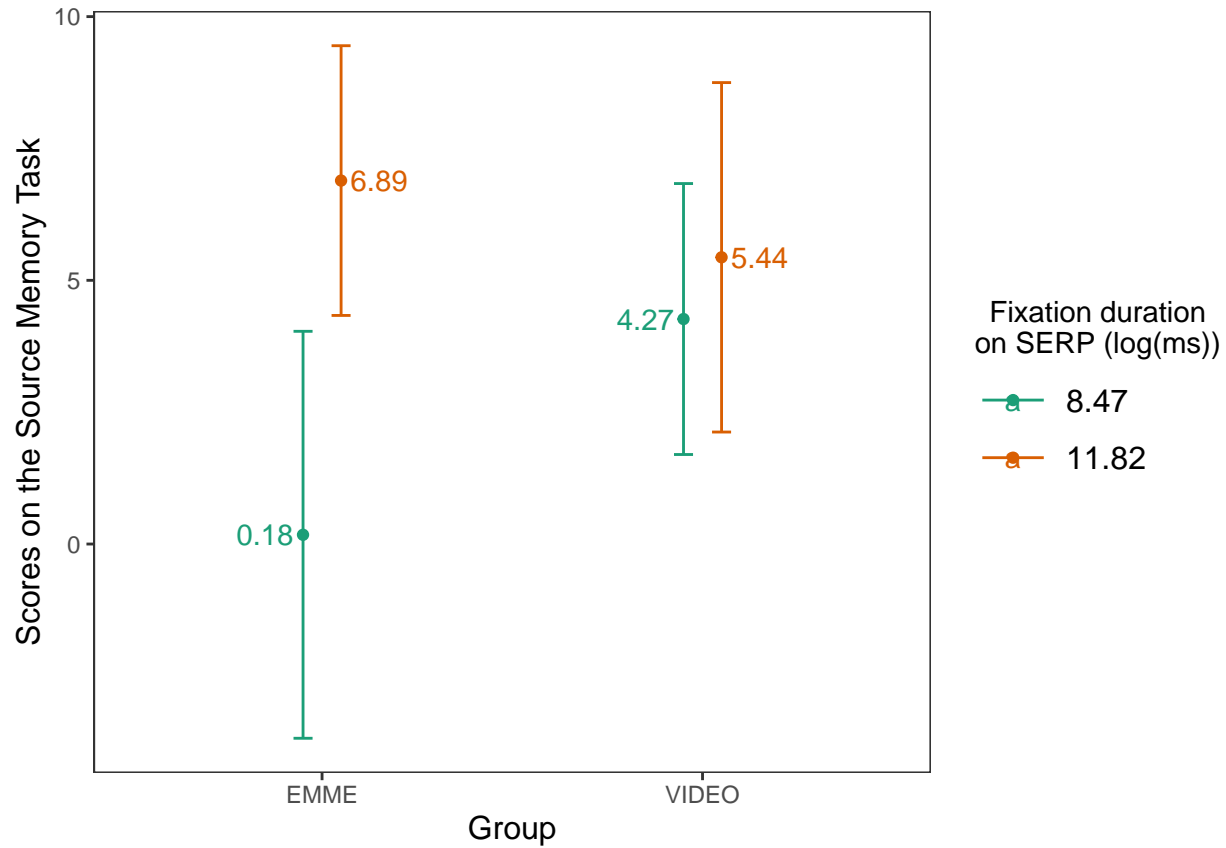


Figure 15. Interaction between Group and Fixation duration on SERP

Figure 16

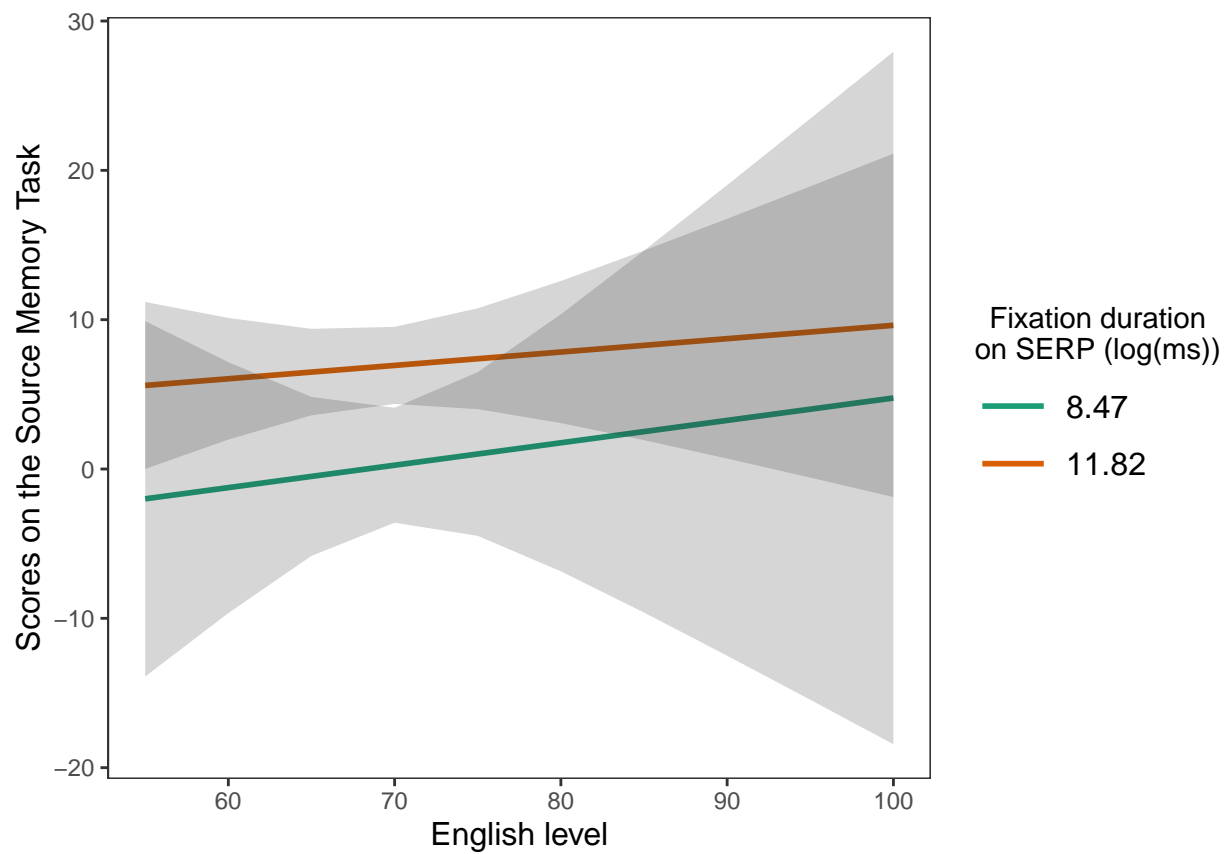


Figure 16. Interaction between English level and Fixation duration on SERP

Figure 17

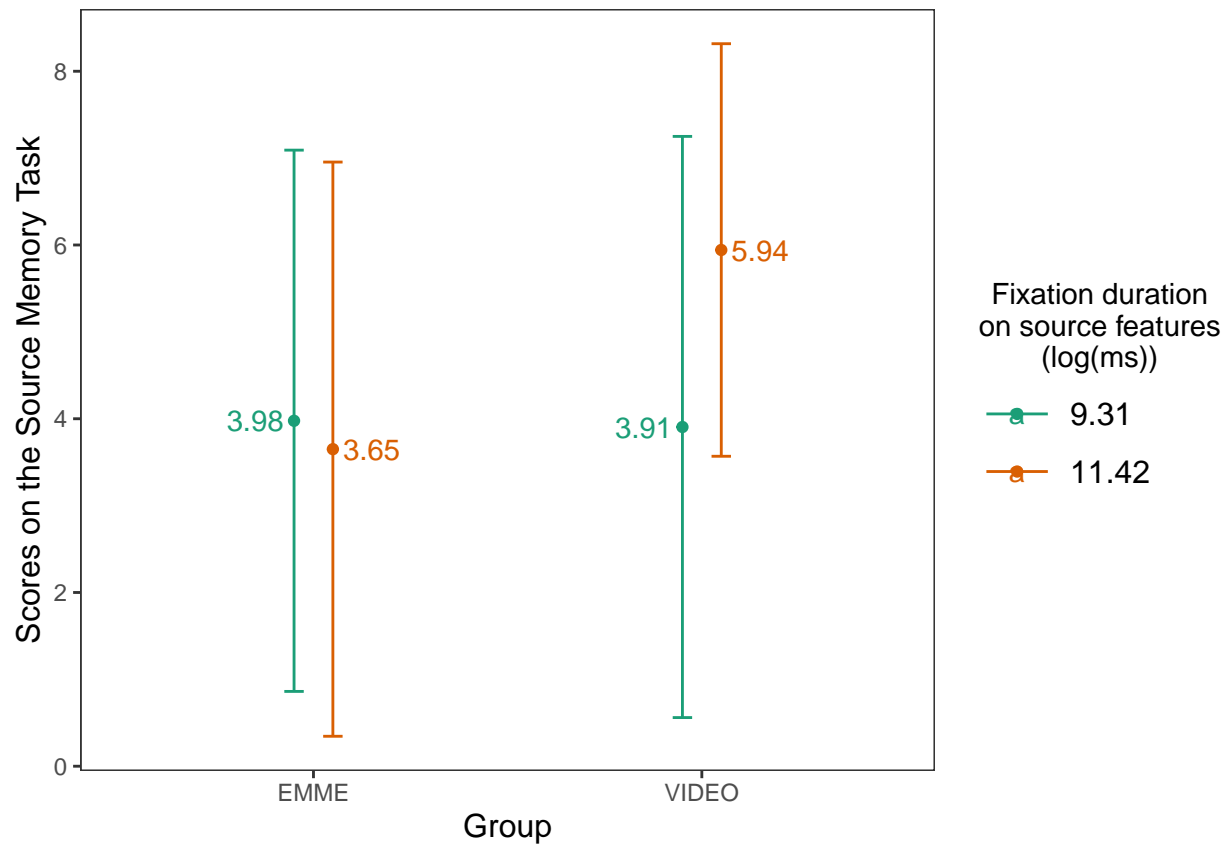


Figure 17. Interaction between Group and Fixation duration on source features

Figure 18

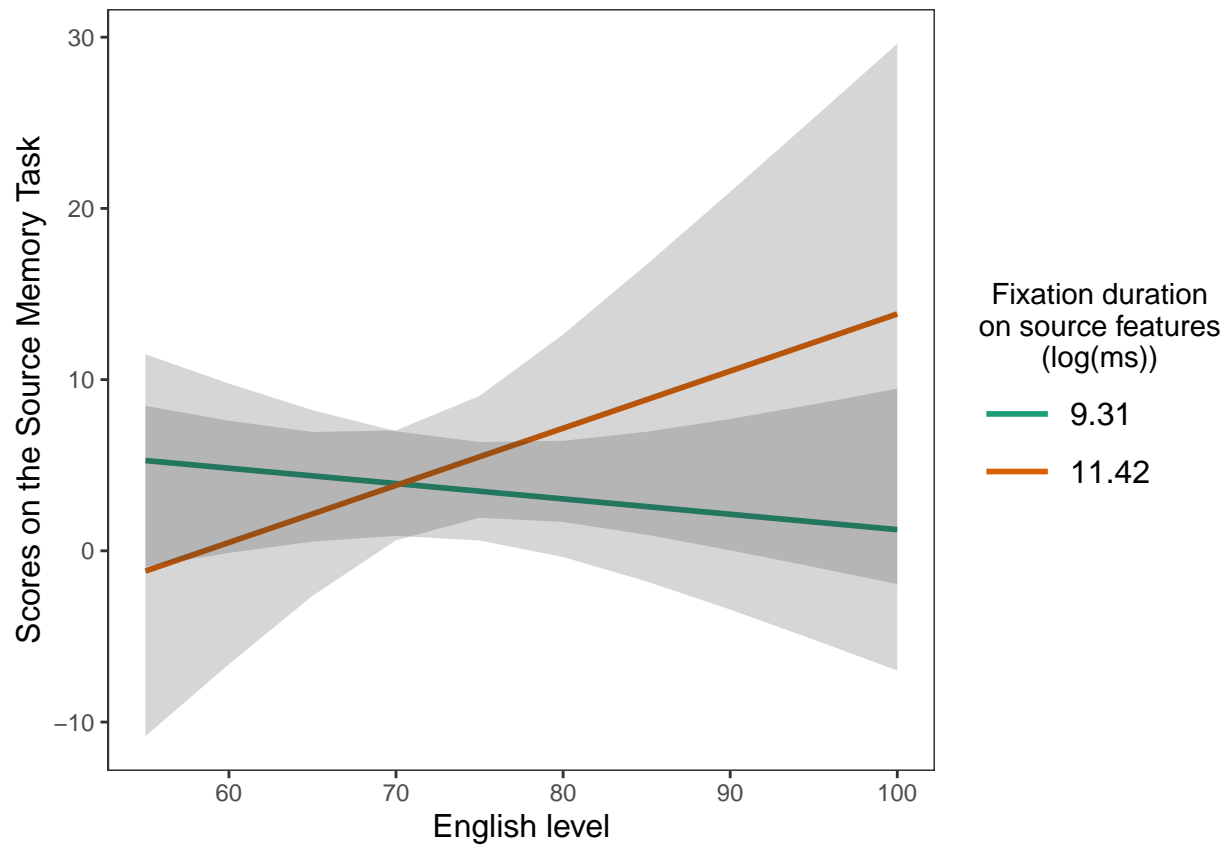


Figure 18. Interaction between English level and Fixation duration on source features

Figure 19

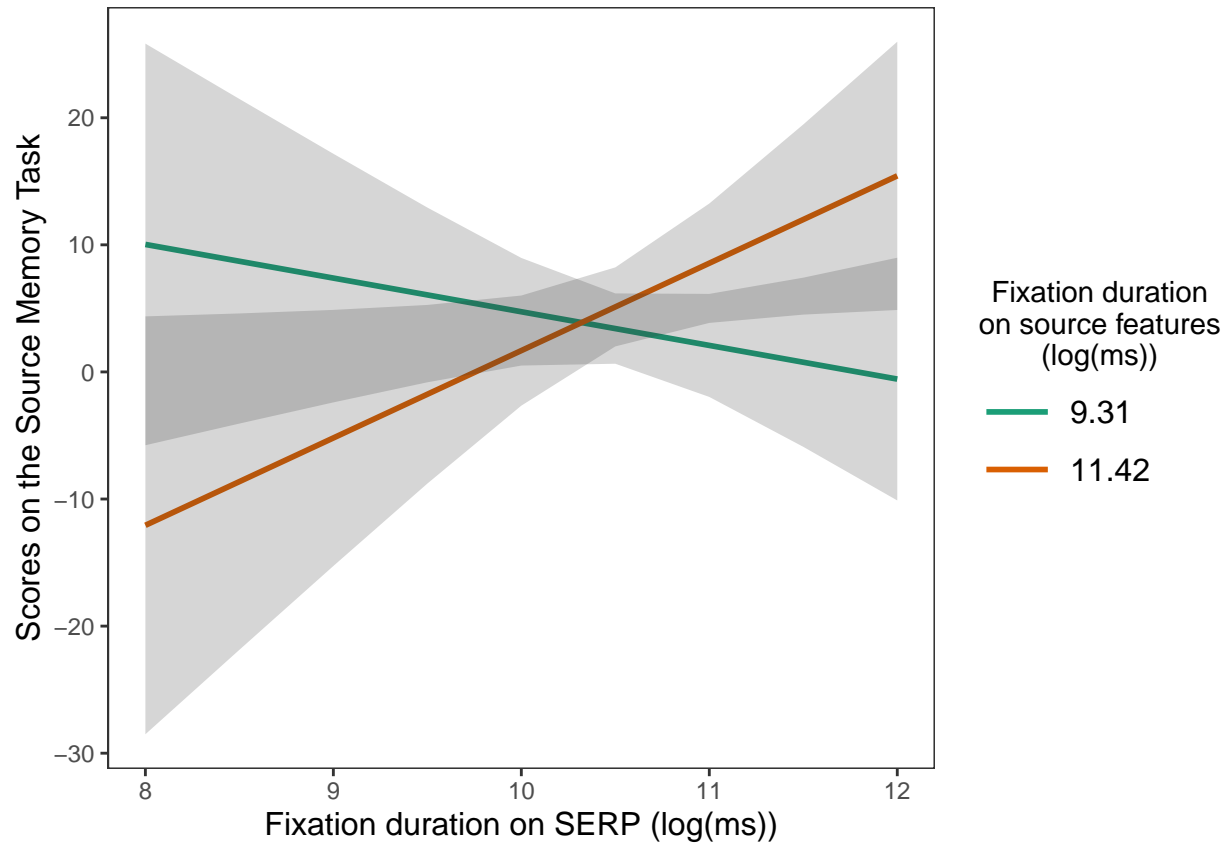


Figure 19. Interaction between Fixation duration on SERP and Fixation duration on source features

Figure 20

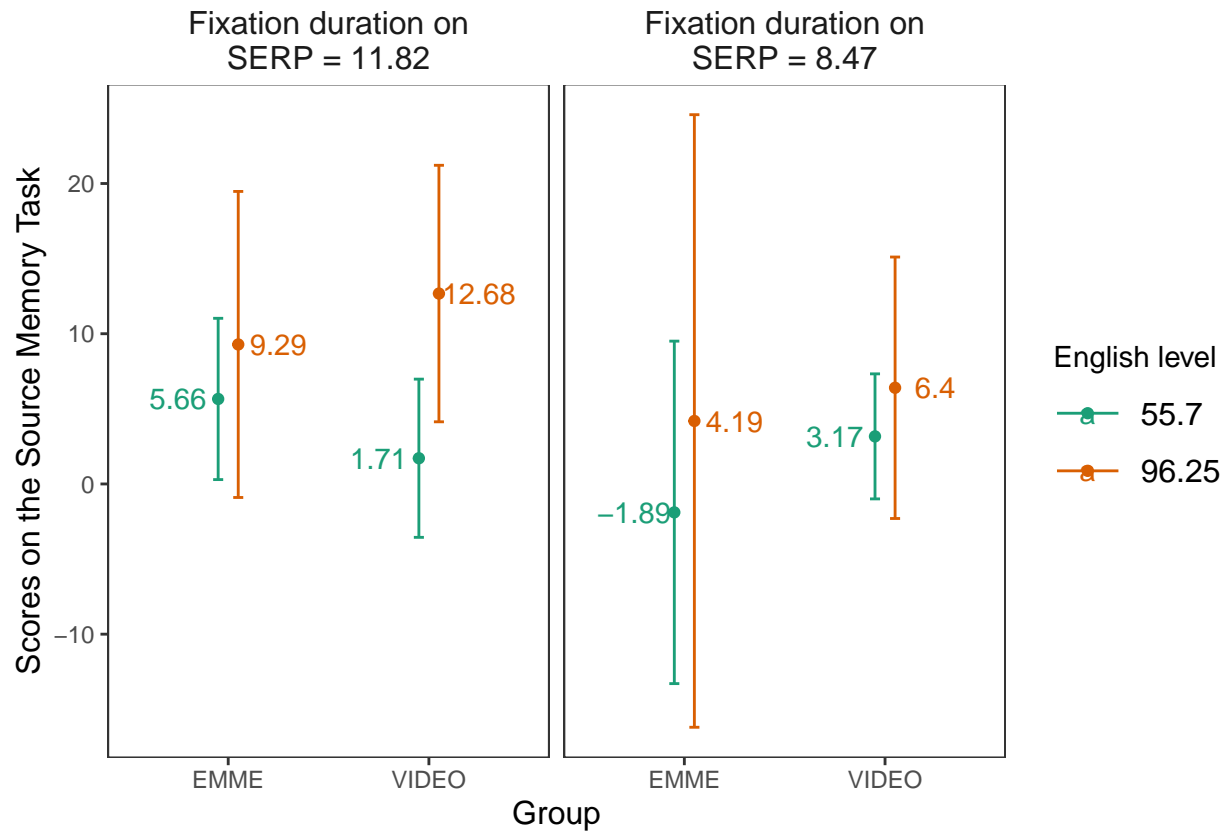


Figure 20. Interaction between Group, English level and Fixation duration on SERP

Figure 21

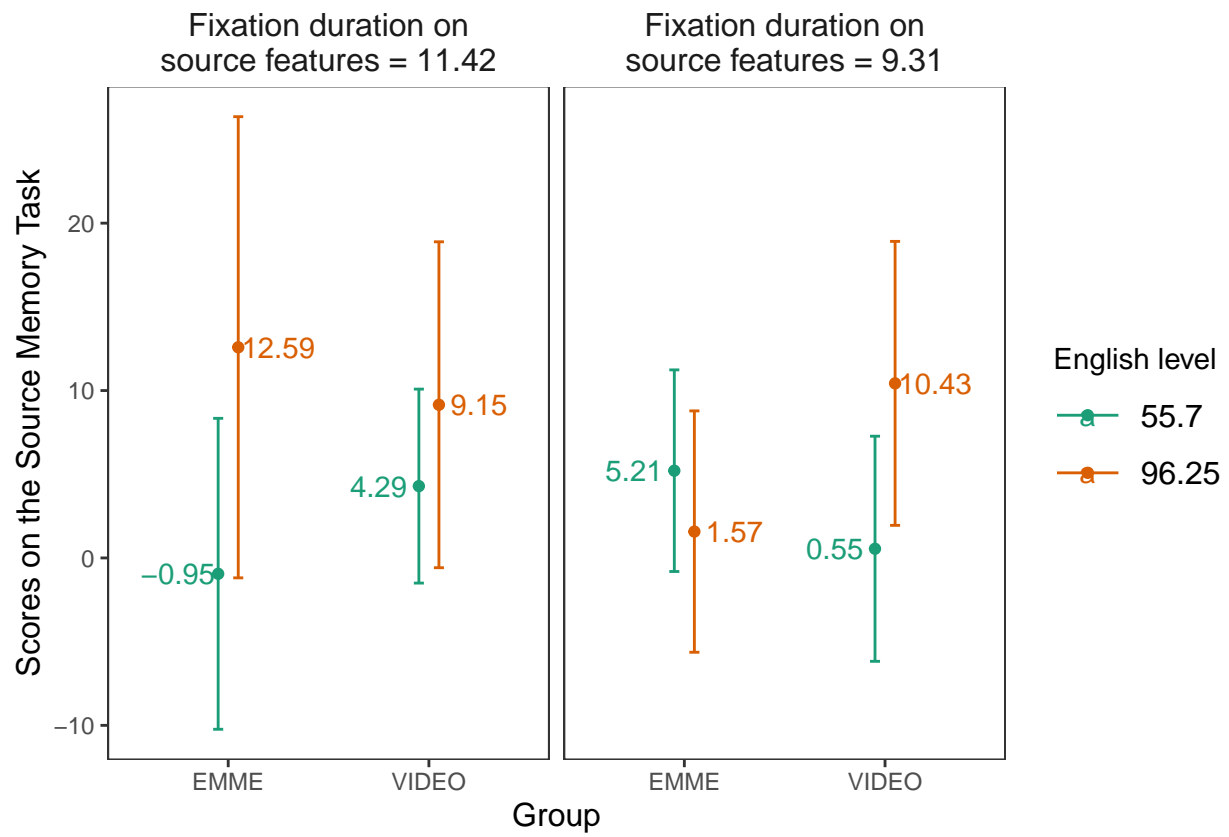


Figure 22

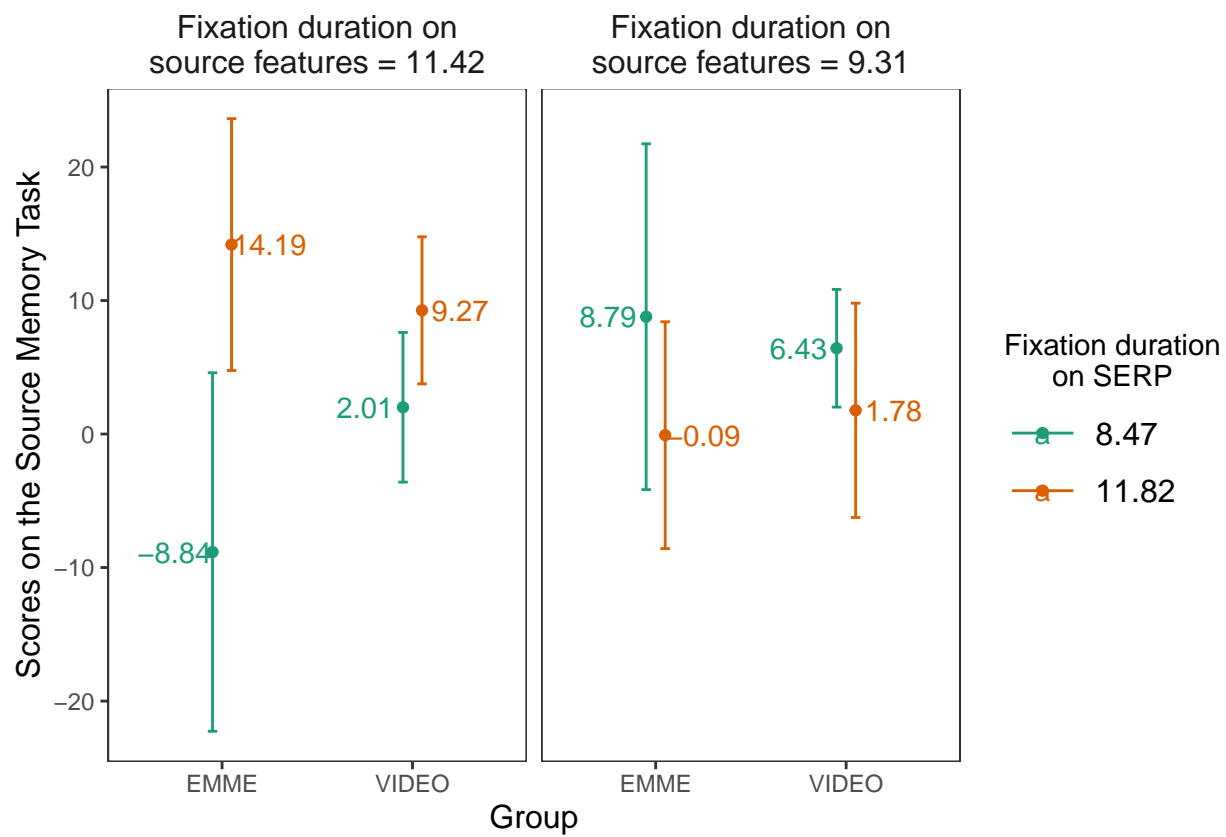
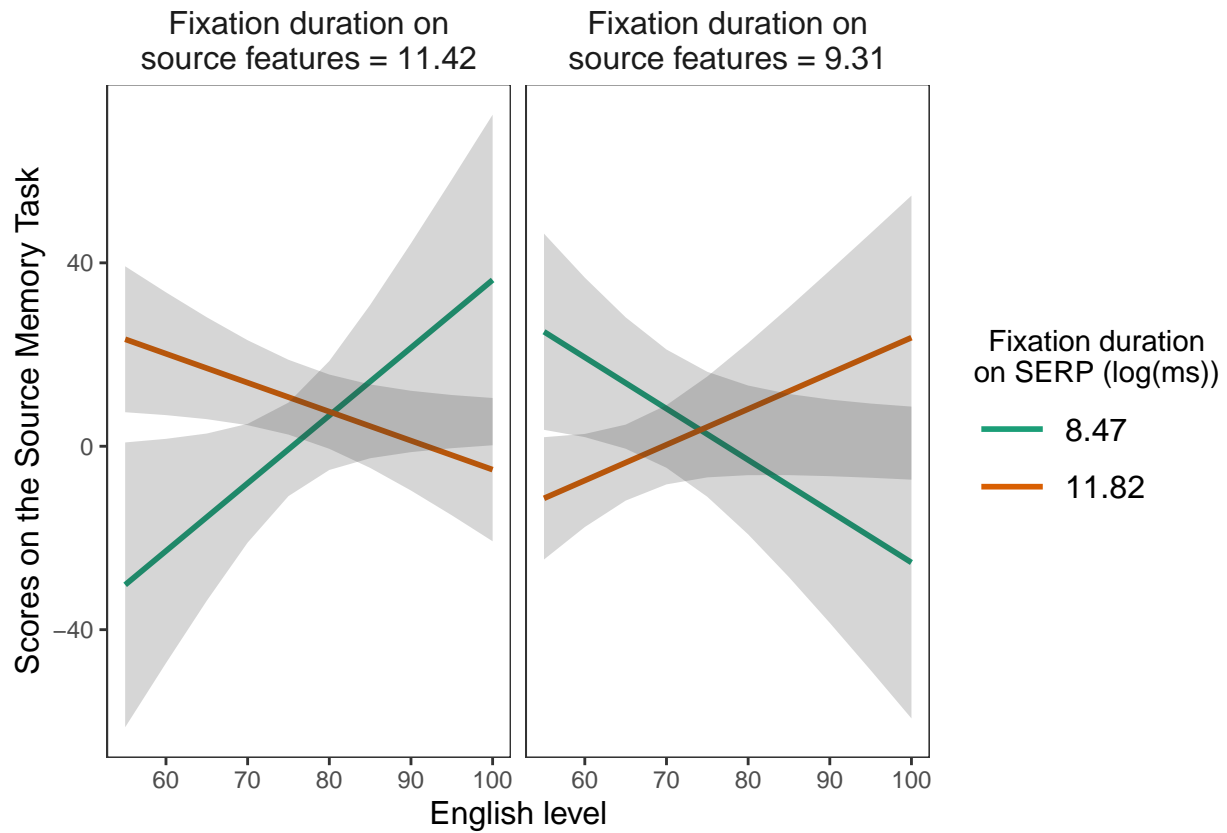


Figure 23



References

SALMERÓN, L.; DELGADO, P.; MASON, L. Using eye-movement modelling examples to improve critical reading of multiple webpages on a conflicting topic. **Journal of Computer Assisted Learning**, vol. 36, no. 6, p. 1038–1051, 2020. <https://doi.org/10.1111/jcal.12458>.