**Appendix A. Generalized Linear Mixed-Effects Modeling**

The R script and the raw data file is available in <https://osf.io/gzamc/?view_only=8be17740e58a4e4a9810fd2caf02bf62>

**1. Fixed and Random Effects Structures in the Maximal Model**

Model building was conducted with the glmmTMB package (v1.1.7; Brooks et al., 2017) in R (v4.3.1; R Core Team, 2023). Specifically, we started with a theoretically driven maximal model (Barr et al., 2013) and empirically modified the random effects structure to address convergence problems (Meteyard & Davies, 2020). The maximal model is presented below:

Posttest Score ~ Group + LE1 + LE2 + LABJ + Group:LE1 + Group:LE2 + Group:LABJ

+ (1 | ID) + (1 + Group + LE1 + LE2 + LABJ + Group:LE1 + Group:LE2

+ Group:LABJ | Item)

*Note*. Posttest Score: score of productive or receptive posttest.

Group: Short-Short, Short-Long, Long-Short, or Long-Long.

LE1 or LE2: learning efficiency of session 1 or 2.

LABJ: test score of language aptitude battery - Japanese.

ID: participant ID.

Item: word.

The hypothesized fixed effects were four different spacing groups, learning efficiency of session 1 and 2, language aptitude test scores, and the interactions between the groups and each of the other predictors. Regarding the random effects, we hypothesized that participants varied in their average posttest scores, and that the average posttest scores and the effects of spacing groups, learning efficiency of session 1 and 2, language aptitude scores, and their interactions varied among words.

**2. Data Preprocessing and Coding**

*LE1, LE2, LABJ: Grand Mean Centering*

For the three continuous variables, namely, LE1, LE2, and LABJ, we conducted grand mean centering for these between-subject variables, following Brauer and Curtin’s (2018) recommendations. Specifically, centered scores (i.e., LE1.c, LE2.c, or LABJ.c) were calculated by using each participant’s score (i.e., LE1, LE2, or LABJ) minus the mean of all participants’ scores (see formulae below).

LE1.c = LE1 – mean of all participants’ LE1

LE2.c = LE2 – mean of all participants’ LE2

LABJ.c = LABJ – mean of all participants’ LABJ

*Group: Deviation Coding*

As for the categorical variable, Group, deviation coding was conducted as recommended by Barr et al. (2013). In order to allow examining pairwise comparison between the four groups, we created four coding schemes for Group, with each group serving as the reference level, respectively. Then we ran the same final mixed-effects model with each coding scheme and summarized the statistically significant pairwise comparisons. Notably, coding schemes affect the results and interpretation of fixed effects and random effects, but not those concerning the whole model - R2 and model diagnostics remain the same regardless of coding schemes.

To facilitate interpreting the results, we followed Field’s (2018) calculation and interpretation of the model coefficients, which is illustrated below in the coding scheme *a) Short-Short as the reference level* (see He & Loewen, 2023, for similar calculations and interpretations).

a) Short-Short as the reference level:

|  |  |  |  |
| --- | --- | --- | --- |
|  ContrastGroup | (2) Short-Long vs.(1) Short-Short | (3) Long-Short vs.(1) Short-Short | (4) Long-Long vs.(1) Short-Short |
| (1) Short-Short | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ |
| (2) Short-Long | $$\frac{3}{4}$$ | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ |
| (3) Long-Short | $$-\frac{1}{4}$$ | $$\frac{3}{4}$$ | $$-\frac{1}{4}$$ |
| (4) Long-Long | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ | $$\frac{3}{4}$$ |

Based on Field (2018), a simplified equation was created to describe the relationship between Posttest Score and Group (Contrast21 = Short-Long vs. Short-Short; Contrast31 = Long-Short vs. Short-Short; Contrast41 = Long-Long vs. Short-Short):

Posttest Score = b0 + b1Contrast21 + b2Contrast31 + b3Contrast41 (a1)

Then for each group, we calculated the average posttest score by replacing Contrast21, Contrast31, and Contrast41 in equation (a1) with the coding values in the table above:

MShort-Short = b0 $- \frac{1}{4}$b1 $- \frac{1}{4}$b2 $- \frac{1}{4}$ b3 (a2)

MShort-Long = b0 $+ \frac{3}{4}$b1 $- \frac{1}{4}$b2 $- \frac{1}{4}$ b3 (a3)

MLong-Short = b0 $- \frac{1}{4}$b1 $+ \frac{3}{4}$b2 $- \frac{1}{4}$ b3 (a4)

MLong-Long = b0 $- \frac{1}{4}$b1 $- \frac{1}{4}$b2 $+ \frac{3}{4}$ b3 (a5)

Using equations (a2) to (a5), we calculated the difference between group means for the pairwise comparisons specified in the table above:

(2) Short-Long vs. (1) Short-Short:

MShort-Long $-$ MShort-Short = (b0 $+ \frac{3}{4}$b1 $- \frac{1}{4}$b2 $- \frac{1}{4}$ b3) $–$ (b0 $- \frac{1}{4}$b1 $- \frac{1}{4}$b2 $- \frac{1}{4}$ b3) = b1

(3) Long-Short vs. (1) Short-Short:

MLong-Short $-$ MShort-Short = (b0 $- \frac{1}{4}$b1 $+ \frac{3}{4}$b2 $- \frac{1}{4}$ b3) $–$ (b0 $- \frac{1}{4}$b1 $- \frac{1}{4}$b2 $- \frac{1}{4}$ b3) = b2

(4) Long-Long vs. (1) Short-Short:

MLong-Long $-$ MShort-Short = (b0 $- \frac{1}{4}$b1 $- \frac{1}{4}$b2 $+ \frac{3}{4}$ b3) $–$ (b0 $- \frac{1}{4}$b1 $- \frac{1}{4}$b2 $- \frac{1}{4}$ b3) = b3

Accordingly, in the current coding scheme, we interpreted b1 as the mean difference between (2) Short-Long and (1) Short-Short, andb2 as the mean difference between (3) Long-Short and (1) Short-Short, and b3 as the mean difference between (4) Long-Long and (1) Short-Short.

Similar calculations and interpretations were adopted for the other three coding schemes. To avoid repetition, in the following, we report the coding table and interpretation without describing the step-by-step equation calculation.

b) Short-Long as the reference level:

|  |  |  |  |
| --- | --- | --- | --- |
|  ContrastGroup | (2) Short-Short vs.(1) Short-Long | (3) Long-Short vs.(1) Short-Long | (4) Long-Long vs.(1) Short-Long |
| (1) Short-Long | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ |
| (2) Short-Short | $$\frac{3}{4}$$ | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ |
| (3) Long-Short | $$-\frac{1}{4}$$ | $$\frac{3}{4}$$ | $$-\frac{1}{4}$$ |
| (4) Long-Long | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ | $$\frac{3}{4}$$ |

(2) Short-Short vs. (1) Short-Long: MShort-Short $-$ MShort-Long = b1

(3) Long-Short vs. (1) Short-Long: MLong-Short $-$ MShort-Long = b2

(4) Long-Long vs. (1) Short-Long: MLong-Long $-$ MShort-Long = b3

c) Long-Short as the reference level:

|  |  |  |  |
| --- | --- | --- | --- |
|  ContrastGroup | (2) Short-Short vs.(1) Long-Short | (3) Short-Long vs.(1) Long-Short | (4) Long-Long vs.(1) Long-Short |
| (1) Long-Short | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ |
| (2) Short-Short | $$\frac{3}{4}$$ | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ |
| (3) Short-Long | $$-\frac{1}{4}$$ | $$\frac{3}{4}$$ | $$-\frac{1}{4}$$ |
| (4) Long-Long | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ | $$\frac{3}{4}$$ |

(2) Short-Short vs. (1) Long-Short: MShort-Short $-$ MLong-Short = b1

(3) Short-Long vs. (1) Long-Short: MShort-Long $-$ MLong-Short = b2

(4) Long-Long vs. (1) Long-Short: MLong-Long $-$ MLong-Short = b3

d) Long-Long as the reference level:

|  |  |  |  |
| --- | --- | --- | --- |
|  ContrastGroup | (2) Short-Short vs.(1) Long-Long | (3) Short-Long vs.(1) Long-Long | (4) Long-Short vs.(1) Long-Long |
| (1) Long-Long | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ |
| (2) Short-Short  | $$\frac{3}{4}$$ | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ |
| (3) Short-Long | $$-\frac{1}{4}$$ | $$\frac{3}{4}$$ | $$-\frac{1}{4}$$ |
| (4) Long-Short | $$-\frac{1}{4}$$ | $$-\frac{1}{4}$$ | $$\frac{3}{4}$$ |

(2) Short-Short vs. (1) Long-Long: MShort-Short $-$ MLong-Long = b1

(3) Short-Long vs. (1) Long-Long: MShort-Long $-$ MLong-Long = b2

(4) Long-Short vs. (1) Long-Long: MLong-Short $-$ MLong-Long = b3

**3. Data Distribution and Model Type**

Item-level scores of the productive and receptive posttests were entered into the mixed-effects models as the outcome variable, respectively. As these scores were 0 or 1, binominal distribution was chosen to build mixed logit models (see Jaeger, 2008).

4. Effect Size and Model Diagnostics

Apart from reporting the results of the final mixed logit models, we calculated the effect size of the whole model and the fixed effects with the performance package (v0.10.4; Lüdecke et al., 2021). Following Gries’ (2021) recommendations to perform model diagnostics, we also checked the residual patterns of the final models with the DHARMa package (v0.4.6; Hartig, 2022).

**References**

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2), 378–400. https://doi.org/10.32614/RJ-2017-066

Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). Los Angeles: SAGE.

Gries, S. T. (2021). (Generalized linear) Mixed‐effects modeling: A learner corpus example. *Language Learning*, *71*(3), 757–798.

Hartig, F. (2022). DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models (version 0.4.6). Retrieved from https://cran.r-project.org/package=DHARMa

He, X., & Loewen, S. (2023). How to present L2 Chinese words effectively for learning: Exploring learning outcomes and learner perceptions. *Studies in Second Language Acquisition*, 1–23.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139.

Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language, 112*, 104092.

R Core Team. (2023). R: A language and environment for statistical computing [software]. Vienna, Austria: R Foundation for Statistical Computing.

**Appendix B. Correlation Coefficients and Scatterplots**

**Entire sample**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Productive test |  | Receptive test |
| 　 | *r* | 95% CI | *p*  | 　 | *r* | 95% CI | *p* |
| Learning Efficiency Session 1 | .00 | [−.15, .15] | .96 | −.07 | [−.24, .09] | .44 |
| Session 2 | −.64 | [−.74, −.53] | < .001 | −.62 | [−.72, −.50] | < .001 |
| LABJ | .21 | [.03, .38] | .02 | .31 | [.14, .48] | < .001 |



**By Group**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Productive test |  | Receptive test |
| 　 | *r* | 95% CI | *p* | 　 | *r* | 95% CI | *p* |
| *Short-Short* |  |  |  |  |  |  |  |
| Learning Efficiency Session 1 | −.42 | [−.65, −.11] | .02 |  | −.61 | [−.81, −.33] | < .001 |
| Session 2 | −.74 | [−.87, −.53] | < .001 |  | −.70 | [−.88, −.49] | < .001 |
| LABJ | .27 | [.00, .50] | .13 |  | .37 | [.02, .64] | .04 |
| *Short-Long* |  |  |  |  |  |  |  |
| Learning Efficiency Session 1 | −.42 | [−.62, −.17] | .02 |  | −.51 | [−.75, −.26] | .00 |
| Session 2 | −.48 | [−.71, −.23] | .01 |  | −.64 | [−.82, −.37] | < .001 |
| LABJ | .17 | [−.23, .48] | .38 |  | .15 | [−.27, .54] | .44 |
| *Long-Short* |  |  |  |  |  |  |  |
| Learning Efficiency Session 1 | −.61 | [−.79, −.39] | < .001 |  | −.45 | [−.77, −.10] | .01 |
| Session 2 | −.82 | [−.89, −.71] | < .001 |  | −.77 | [−.89, −.55] | < .001 |
| LABJ | .24 | [−.16, .61] | .20 |  | .42 | [.06, .70] | .02 |
| *Long-Long* |  |  |  |  |  |  |  |
| Learning Efficiency Session 1 | −.46 | [−.71, −.13] | .02 |  | −.54 | [−.74, −.20] | .01 |
| Session 2 | −.79 | [−.93, −.56] | < .001 |  | −.75 | [−.89, −.46] | < .001 |
| LABJ | .44 | [.01, .72] | .03 | 　 | .66 | [.32, .85] | < .001 |

*Short-Short*



*Short-Long*



*Long-Short*



*Long-Long*



**Appendix C. Results of the Final Mixed Logit Model for the Productive Posttest**

**1. Results That Were Shared Across All Coding Schemes**

*Formula:*

Posttest Score ~ Group + LE1 + LE2 + LABJ + Group:LE1 + Group:LE2 + Group:LABJ

+ (1 | ID) + (1 | Item)

*Note*. Group: Short-Short, Short-Long, Long-Short, Long-Long.

LE1 or LE2: learning efficiency of session 1 or 2.

LABJ: test score of language aptitude battery - Japanese.

ID: participant ID.

Item: word.

*Effect Size:*

Conditional (total) *R2*: 0.402

Marginal (fixed effects) *R2*: 0.244

*Model Diagnostics:*

The residual plots below did not suggest serious issues in the final model.



**2. Results of Each Coding Scheme**

**a) Short-Short as the reference level:**

Random Effects:

*Groups Name Variance Std.Dev.*

 ID (Intercept) 0.1996 0.4467

 Item (Intercept) 0.6711 0.8192

Fixed Effects:

 *Estimate Std. Error z value Pr(>|z|)*

(Intercept) -1.244837 0.231742 -5.372 7.8e-08 \*\*\*

Group.SL-SS 0.605146 0.427204 1.417 0.1566

Group.LS-SS 0.468496 0.356422 1.314 0.1887

Group.LL-SS 1.545477 0.290038 5.329 9.9e-08 \*\*\*

LE1 -0.001726 0.009214 -0.187 0.8514

LE2 -0.123332 0.013709 -8.996 < 2e-16 \*\*\*

LABJ -0.015558 0.019690 -0.790 0.4294

Group.SL-SS:LE1 -0.071272 0.029657 -2.403 0.0163 \*

Group.SL-SS:LE2 0.099482 0.041118 2.419 0.0155 \*

Group.SL-SS:LABJ -0.057028 0.061436 -0.928 0.3533

Group.LS-SS:LE1 -0.049518 0.025650 -1.931 0.0535

Group.LS-SS:LE2 0.047490 0.037818 1.256 0.2092

Group.LS-SS:LABJ -0.061497 0.060355 -1.019 0.3082

Group.LL-SS:LE1 -0.016873 0.027370 -0.616 0.5376

Group.LL-SS:LE2 0.064696 0.040468 1.599 0.1099

Group.LL-SS:LABJ -0.019605 0.060284 -0.325 0.7450

*Note*. \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

Group.SL-SS: Short-Long vs. Short-Short.

Group.LS-SS: Long-Short vs. Short-Short.

Group.LL-SS: Long-Long vs. Short-Short.

*----------------Annotations----------------*

Significant Interactions:

Group.SL-SS:LE1 = (MShort-Long $-$ MShort-Short)\*LE1

Group.SL-SS:LE2 = (MShort-Long $-$ MShort-Short)\*LE2

**b) Short-Long as the reference level:**

Random Effects:

*Groups Name Variance Std.Dev.*

 ID (Intercept) 0.1996 0.4467

 Item (Intercept) 0.6711 0.8192

Fixed Effects:

 *Estimate Std. Error z value Pr(>|z|)*

(Intercept) -1.244843 0.231744 -5.372 7.8e-08 \*\*\*

Group.SS-SL -0.605120 0.427203 -1.416 0.1566

Group.LS-SL -0.136633 0.480660 -0.284 0.7762

Group.LL-SL 0.940355 0.433425 2.170 0.0300 \*

LE1 -0.001725 0.009214 -0.187 0.8514

LE2 -0.123332 0.013709 -8.996 < 2e-16 \*\*\*

LABJ -0.015558 0.019690 -0.790 0.4294

Group.SS-SL:LE1 0.071271 0.029656 2.403 0.0163 \*

Group.SS-SL:LE2 -0.099482 0.041118 -2.419 0.0155 \*

Group.SS-SL:LABJ 0.057027 0.061436 0.928 0.3533

Group.LS-SL:LE1 0.021755 0.024679 0.882 0.3780

Group.LS-SL:LE2 -0.051993 0.035824 -1.451 0.1467

Group.LS-SL:LABJ -0.004468 0.050615 -0.088 0.9297

Group.LL-SL:LE1 0.054399 0.026510 2.052 0.0402 \*

Group.LL-SL:LE2 -0.034788 0.038607 -0.901 0.3675

Group.LL-SL:LABJ 0.037422 0.050605 0.740 0.4596

*Note*. \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

Group.SS-SL: Short-Short vs. Short-Long.

Group.LS-SL: Long-Short vs. Short-Long.

Group.LL-SL: Long-Long vs. Short-Long.

*----------------Annotations----------------*

Significant Interactions:

Group.SS-SL:LE1 = (MShort-Short $-$ MShort-Long)\*LE1 (found in the results of Short-Short)

Group.SS-SL:LE2 = (MShort-Short $-$ MShort-Long)\*LE2 (found in the results of Short-Short)

Group.LL-SL:LE1 = (MLong-Long $-$ MShort-Long)\*LE1

**c) Long-Short as the reference level:**

Random Effects:

*Groups Name Variance Std.Dev.*

 ID (Intercept) 0.1996 0.4467

 Item (Intercept) 0.6711 0.8192

Fixed Effects:

 *Estimate Std. Error z value Pr(>|z|)*

(Intercept) -1.244829 0.231744 -5.372 7.81e-08 \*\*\*

Group.SS-LS -0.468475 0.356422 -1.314 0.18872

Group.SL-LS 0.136641 0.480662 0.284 0.77620

Group.LL-LS 1.076995 0.363971 2.959 0.00309 \*\*

LE1 -0.001725 0.009214 -0.187 0.85147

LE2 -0.123332 0.013709 -8.996 < 2e-16 \*\*\*

LABJ -0.015558 0.019690 -0.790 0.42944

Group.SS-LS:LE1 0.049518 0.025650 1.931 0.05354

Group.SS-LS:LE2 -0.047489 0.037818 -1.256 0.20921

Group.SS-LS:LABJ 0.061497 0.060355 1.019 0.30824

Group.SL-LS:LE1 -0.021755 0.024679 -0.882 0.37804

Group.SL-LS:LE2 0.051995 0.035824 1.451 0.14668

Group.SL-LS:LABJ 0.004468 0.050615 0.088 0.92965

Group.LL-LS:LE1 0.032644 0.021936 1.488 0.13671

Group.LL-LS:LE2 0.017206 0.035074 0.491 0.62374

Group.LL-LS:LABJ 0.041891 0.049287 0.850 0.39536

*Note*. \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

Group.SS-LS: Short-Short vs. Long-Short.

Group.SL-LS: Short-Long vs. Long-Short.

Group.LL-LS: Long-Long vs. Long-Short.

**d) Long-Long as the reference level:**

Random Effects:

*Groups Name Variance Std.Dev.*

 ID (Intercept) 0.1996 0.4467

 Item (Intercept) 0.6711 0.8192

Fixed Effects:

 *Estimate Std. Error z value Pr(>|z|)*

(Intercept) -1.244838 0.231744 -5.372 7.8e-08 \*\*\*

Group.SS-LL -1.545478 0.290038 -5.329 9.9e-08 \*\*\*

Group.SL-LL -0.940368 0.433427 -2.170 0.03004 \*

Group.LS-LL -1.077004 0.363972 -2.959 0.00309 \*\*

LE1 -0.001726 0.009214 -0.187 0.85140

LE2 -0.123332 0.013709 -8.996 < 2e-16 \*\*\*

LABJ -0.015557 0.019690 -0.790 0.42946

Group.SS-LL:LE1 0.016873 0.027370 0.616 0.53758

Group.SS-LL:LE2 -0.064694 0.040468 -1.599 0.10990

Group.SS-LL:LABJ 0.019607 0.060284 0.325 0.74500

Group.SL-LL:LE1 -0.054399 0.026511 -2.052 0.04017 \*

Group.SL-LL:LE2 0.034789 0.038607 0.901 0.36754

Group.SL-LL:LABJ -0.037422 0.050605 -0.739 0.45961

Group.LS-LL:LE1 -0.032644 0.021936 -1.488 0.13672

Group.LS-LL:LE2 -0.017206 0.035075 -0.491 0.62374

Group.LS-LL:LABJ -0.041890 0.049287 -0.850 0.39537

*Note*. \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

Group.SS-LL: Short-Short vs. Long-Long.

Group.SL-LL: Short-Long vs. Long-Long.

Group.LS-LL: Long-Short vs. Long-Long.

*----------------Annotations----------------*

Significant Interaction:

Group.SL-LL:LE1 = (MShort-Long $-$ MLong-Long)\*LE1 (found in the results of Short-Long)

**3. Summary of Significant Interactions**

1) (MShort-Long $-$ MShort-Short)\*LE1

2) (MShort-Long $-$ MLong-Long)\*LE1

3) (MShort-Long $-$ MShort-Short)\*LE2

**Appendix D. Results of the Final Mixed Logit Model for the Receptive Posttest**

**1. Results That were Shared Across All Coding Schemes**

*Formula:*

Posttest Score ~ Group + LE1 + LE2 + LABJ + Group:LE1 + Group:LE2 + Group:LABJ

+ (1 | ID) + (1 | Item)

*Note*. Group: Short-Short, Short-Long, Long-Short, Long-Long.

LE1 or LE2: learning efficiency of session 1 or 2.

LABJ: test score of language aptitude battery - Japanese.

ID: participant ID.

Item: word.

*Effect Size:*

Conditional (total) *R2*: 0.439

Marginal (fixed effects) *R2*: 0.194

*Model Diagnostics:*

The residual plots below did not reveal serious issues in the final model.



**2. Results of Each Coding Scheme**

**a) Short-Short as the reference level:**

Random Effects:

*Groups Name Variance Std.Dev.*

 ID (Intercept) 0.3064 0.5536

 Item (Intercept) 1.1367 1.0661

Fixed Effects:

 *Estimate Std. Error z value Pr(>|z|)*

(Intercept) 0.475249 0.281191 1.690 0.09100

Group.SL-SS 1.251126 0.437258 2.861 0.00422 \*\*

Group.LS-SS 0.962132 0.364782 2.638 0.00835 \*\*

Group.LL-SS 1.783309 0.326132 5.468 4.55e-08 \*\*\*

LE1 -0.022335 0.010066 -2.219 0.02650 \*

LE2 -0.094393 0.014361 -6.573 4.94e-11 \*\*\*

LABJ 0.006291 0.019815 0.318 0.75085

Group.SL-SS:LE1 -0.006094 0.030339 -0.201 0.84081

Group.SL-SS:LE2 -0.025262 0.040107 -0.630 0.52878

Group.SL-SS:LABJ -0.089516 0.056769 -1.577 0.11483

Group.LS-SS:LE1 0.025666 0.027629 0.929 0.35291

Group.LS-SS:LE2 -0.017997 0.037203 -0.484 0.62856

Group.LS-SS:LABJ -0.020017 0.058056 -0.345 0.73025

Group.LL-SS:LE1 0.030745 0.031032 0.991 0.32180

Group.LL-SS:LE2 -0.009143 0.043146 -0.212 0.83218

Group.LL-SS:LABJ 0.071208 0.059299 1.201 0.22981

*Note*. \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

Group.SL-SS: Short-Long vs. Short-Short.

Group.LS-SS: Long-Short vs. Short-Short.

Group.LL-SS: Long-Long vs. Short-Short.

**b) Short-Long as the reference level:**

Random Effects:

*Groups Name Variance Std.Dev.*

 ID (Intercept) 0.3064 0.5536

 Item (Intercept) 1.1367 1.0662

Fixed Effects:

 *Estimate Std. Error z value Pr(>|z|)*

(Intercept) 0.475245 0.281191 1.690 0.09101

Group.SS-SL -1.251130 0.437260 -2.861 0.00422 \*\*

Group.LS-SL -0.289001 0.498606 -0.580 0.56217

Group.LL-SL 0.532175 0.469679 1.133 0.25719

LE1 -0.022335 0.010066 -2.219 0.02650 \*

LE2 -0.094393 0.014361 -6.573 4.94e-11 \*\*\*

LABJ 0.006290 0.019815 0.317 0.75092

Group.SS-SL:LE1 0.006092 0.030339 0.201 0.84087

Group.SS-SL:LE2 0.025262 0.040107 0.630 0.52878

Group.SS-SL:LABJ 0.089514 0.056769 1.577 0.11484

Group.LS-SL:LE1 0.031758 0.025552 1.243 0.21391

Group.LS-SL:LE2 0.007262 0.037183 0.195 0.84515

Group.LS-SL:LABJ 0.069494 0.052616 1.321 0.18657

Group.LL-SL:LE1 0.036839 0.029218 1.261 0.20737

Group.LL-SL:LE2 0.016118 0.043146 0.374 0.70873

Group.LL-SL:LABJ 0.160724 0.054055 2.973 0.00295 \*\*

*Note*. \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

Group.SS-SL: Short-Short vs. Short-Long.

Group.LS-SL: Long-Short vs. Short-Long.

Group.LL-SL: Long-Long vs. Short-Long.

*----------------Annotations----------------*

Significant Interaction:

Group.LL-SL:LABJ = (MLong-Long $-$ MShort-Long)\*LABJ

**c) Long-Short as the reference level:**

Random Effects:

 *Groups Name Variance Std.Dev.*

 ID (Intercept) 0.3064 0.5536

 Item (Intercept) 1.1366 1.0661

Fixed Effects:

 *Estimate Std. Error z value Pr(>|z|)*

(Intercept) 0.475251 0.281188 1.690 0.09100

Group.SS-LS -0.962128 0.364783 -2.638 0.00835 \*\*

Group.SL-LS 0.288992 0.498605 0.580 0.56218

Group.LL-LS 0.821181 0.404189 2.032 0.04219 \*

LE1 -0.022335 0.010066 -2.219 0.02650 \*

LE2 -0.094393 0.014361 -6.573 4.94e-11 \*\*\*

LABJ 0.006292 0.019815 0.318 0.75085

Group.SS-LS:LE1 -0.025665 0.027629 -0.929 0.35293

Group.SS-LS:LE2 0.017998 0.037203 0.484 0.62855

Group.SS-LS:LABJ 0.020018 0.058057 0.345 0.73024

Group.SL-LS:LE1 -0.031759 0.025552 -1.243 0.21390

Group.SL-LS:LE2 -0.007262 0.037183 -0.195 0.84515

Group.SL-LS:LABJ -0.069496 0.052616 -1.321 0.18657

Group.LL-LS:LE1 0.005080 0.026369 0.193 0.84724

Group.LL-LS:LE2 0.008853 0.040438 0.219 0.82671

Group.LL-LS:LABJ 0.091226 0.055373 1.647 0.09946

*Note*. \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

Group.SS-LS: Short-Short vs. Long-Short.

Group.SL-LS: Short-Long vs. Long-Short.

Group.LL-LS: Long-Long vs. Long-Short.

**d) Long-Long as the reference level:**

Random Effects:

*Groups Name Variance Std.Dev.*

 ID (Intercept) 0.3064 0.5536

 Item (Intercept) 1.1367 1.0661

Fixed Effects:

 *Estimate Std. Error z value Pr(>|z|)*

(Intercept) 0.475238 0.281189 1.690 0.09101

Group.SS-LL -1.783306 0.326132 -5.468 4.55e-08 \*\*\*

Group.SL-LL -0.532180 0.469679 -1.133 0.25718

Group.LS-LL -0.821198 0.404188 -2.032 0.04218 \*

LE1 -0.022335 0.010066 -2.219 0.02650 \*

LE2 -0.094393 0.014361 -6.573 4.94e-11 \*\*\*

LABJ 0.006290 0.019815 0.317 0.75090

Group.SS-LL:LE1 -0.030745 0.031032 -0.991 0.32181

Group.SS-LL:LE2 0.009143 0.043146 0.212 0.83217

Group.SS-LL:LABJ -0.071208 0.059299 -1.201 0.22982

Group.SL-LL:LE1 -0.036837 0.029218 -1.261 0.20740

Group.SL-LL:LE2 -0.016117 0.043146 -0.374 0.70875

Group.SL-LL:LABJ -0.160721 0.054055 -2.973 0.00295 \*\*

Group.LS-LL:LE1 -0.005078 0.026369 -0.193 0.84730

Group.LS-LL:LE2 -0.008855 0.040438 -0.219 0.82667

Group.LS-LL:LABJ -0.091223 0.055373 -1.647 0.09947

*Note*. \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

Group.SS-LL: Short-Short vs. Long-Long.

Group.SL-LL: Short-Long vs. Long-Long.

Group.LS-LL: Long-Short vs. Long-Long.

*----------------Annotations----------------*

Significant Interaction:

Group.SL-LL:LABJ = (MShort-Long $- $MLong-Long)\*LABJ (found in the results of Short-Long)

**3. Summary of Significant Interaction**

1) (MLong-Long $-$ MShort-Long)\*LABJ