Online Supplementary Materials

Appendices

**Appendix A**

***Proficiency Measures***

The cloze test was a multiple-choice test, including ten items. Each item had three options. The test was based on a 162-word text, which was taken from a booklet, including a mock examination for the B1 level TELC Spanish Proficiency Exam (downloadable from https://www.telc.net/en/candidates/language-examinations/tests/detail/telc-espanol-b1.html#t=2). Learners were awarded one point for each correct response.

The oral picture description task was based on a picture taken from a spot-the-difference task from New English File Intermediate by Oxford University Press. The picture depicts a busy street full of people doing various things. Learners’ descriptions were audio-recorded and then rated by two native speakers of Spanish (using a scale of 0 to 3) on four categories: 1) coherence; 2) range; 3) fluency; 4) accuracy. Then, an overall oral proficiency score was calculated for every participant by averaging the four scores.

*Descriptive Statistics for Tests Measuring Proficiency*

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Cloze | Oral Picture Description  |
|  | *N* | *M* | *SD* | 95% CI | *M* | *SD* | 95% CI |
| Control  | 42 | 0.61 | 0.42 | 0.55, 0.66 | 1.91 | 0.69 | 1.69, 2.12 |
| Implicit | 35 | 0.58 | 0.19 | 0.51, 0.65 | 1.83 | 0.89 | 1.52, 2.13 |
| Explicit | 35 | 0.54 | 0.2 | 0.48, 0.61 | 1.72 | 0.78 | 1.46, 1.49 |

**Appendix B**

***Descriptive and Inferential Statistics for Number of Feedback Instances***

*Descriptive Statistics for Number of Feedback Instances Per Task*

|  |  |  |
| --- | --- | --- |
|  | Gender | DOM |
|  | Task 1 | Task 2 | Both Tasks  | Task 1 | Task 2 | Both Tasks  |
|  | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Implicit | 6.44 | 2.74 | 5.28 | 3.46 | 11.72 | 5.45 | 11.39 | 4.59 | 8.17 | 4.54 | 19.56 | 8.11 |
| Explicit | 7.63 | 3.48 | 5.97 | 3.61 | 13.43 | 6.25 | 9.51 | 3.82 | 5.17 | 2.91 | 14.58 | 5.38 |

***Inferential Statistics for Number of Feedback Instances***

*Analysis 1*

We statistically controlled the amount of feedback using two custom ANCOVA models, one for each gain score type (GJT DOM and OPT DOM). We entered feedback amount, Explicit Learning Ability, Implicit Memory Ability, Implicit Learning Ability as covariates in the model. The models also included the two-way interaction terms between the covariates and Group. We did not include the control group in these analyses because this group did not receive feedback. As reported in the main text, the results of the custom ANCOVA on DOM scores which did not include feedback amount as a covariate have indicated that the interaction effect between Explicit Learning Ability and Group was statistically significant. The results summarized in the table below also show that the interaction between Group and Explicit Learning Ability is still significant even after controlling for feedback amount.

*Custom ANCOVA Models with Feedback Amount as a Covariate*

|  |  |  |
| --- | --- | --- |
|  | OPT | GJT |
| Factors | *df* | *F* | *p* | *η2* | *df* | *F* | *p* | *η2* |
| group | 1 | 3.28 | 0.075 | 0.040 | 1 | 0.752 | 0.389 | 0.011 |
| explicit learning | 1 | 0.261 | 0.612 | 0.003 | 1 | 0.222 | 0.639 | 0.003 |
| implicit memory | 1 | 0.488 | 0.488 | 0.006 | 1 | 1.227 | 0.272 | 0.018 |
| implicit learning | 1 | 0.036 | 0.849 | 0.000 | 1 | 0.047 | 0.829 | 0.001 |
| Feedback Amount | 1 | 8.086 | 0.006 | 0.099 | 1 | 3.395 | 0.070 | 0.050 |
| group \* explicit learning | 1 | 7.302 | 0.009 | 0.090 | 1 | 0.220 | 0.641 | 0.003 |
| group \* implicit memory | 1 | 0.269 | 0.606 | 0.003 | 1 | 0.063 | 0.803 | 0.001 |
| group \* implicit learning  | 1 | 0.056 | 0.814 | 0.000 | 1 | 0.008 | 0.927 | 0.000 |
| Error | 58 |  |  |  | 59 |   |   |   |

*Analysis 2*

 The use of ANCOVA is not recommended to analyze designs in which (Keppel & Wickens, 2004) the covariate is affected by the treatment. One may argue that the differences in feedback amount between the feedback groups were the outcome of the treatment, which was feedback type in the current study. When an ANCOVA is used in such cases, it can remove a portion of the treatment effect. Huitema (2011) recommends using a quasi-ANCOVA procedure, which “eliminates the between-groups adjustment that causes bias (when treatments affect covariates) but retains the within-group adjustment that increases precision” (Huitema, 2011, p. 308). This analysis is carried out by substituting residuals of ANOVA performed on the covariates for the original covariates. We carried out two quasi-ANCOVAs, one for each score type, to determine whether the inclusion of feedback amount as a covariate changes the interaction effect between the cognitive aptitudes and Group. Feedback amount, Explicit Learning Ability, Implicit Memory Ability, Implicit Learning Ability were the covariates in these models. The models also included the interaction terms between the covariates and Group. The control group was not included in these models since they did not receive any feedback. The results showed that the inclusion of feedback amount did not change the results of the custom made ANCOVAs that did not include feedback amount as a covariate. The interaction between Explicit Learning Ability and Group in the model for the OPT scores was still significant.

*Custom Quasi-ANCOVA Models with Feedback Amount as a Covariate*

|  |  |  |
| --- | --- | --- |
|  | GJT | OPT |
|  | *df* | *F* | *p* | *η*2 | *df* | *F* | *p* | *η*2 |
| Group | 1 | 2.604 | 0.112 | 0.038 | 1 | 8.871 | 0.004 | 0.109 |
| Explicit Learning | 1 | 0.222 | 0.639 | 0.003 | 1 | 0.261 | 0.612 | 0.003 |
| Implicit Memory | 1 | 1.227 | 0.272 | 0.018 | 1 | 0.488 | 0.488 | 0.006 |
| Implicit Learning | 1 | 0.047 | 0.829 | 0.001 | 1 | 0.036 | 0.849 | 0.000 |
| Feedback Amount | 1 | 3.395 | 0.07 | 0.050 | 1 | 8.086 | 0.006 | 0.099 |
| Group \* Explicit Learning | 1 | 0.220 | 0.641 | 0.003 | 1 | 7.302 | 0.009 | 0.090 |
| group \* Implicit Memory | 1 | 0.063 | 0.803 | 0.001 | 1 | 0.269 | 0.606 | 0.003 |
| Group \* Implicit Learning  | 1 | 0.008 | 0.927 | 0.000 | 1 | 0.056 | 0.814 | 0.000 |
| Error | 59 |  |  |  | 58 |  |  |  |

*References Used for the Analyses*

Huitema, B. (2011). *The analysis of covariance and alternatives: statistical methods for experiments, quasi-experiments, and single-case studies*. Hoboken, NJ: Wiley.

Keppel, G. & Wickens, T.D. (2004). *Design and analysis*: *A* *researcher’s handbook*. *4th edition*. Upper Saddle River, NJ: Pearson.

**Appendix C**

***Determining the Cutoff Score***

We used the following procedures to determine the cutoff score of 0.80. We first calculated the average gain made across groups and outcome measures, which was 0.15 (*SD* = .22). We then added 0.05 to this number to make the cutoff more conservative and deducted the sum of these two numbers (0.15 + 0.5 = 0.20) from the ceiling (1.0).

Four learners scored above 0.80 in the GJT on ungrammatical gender items. No cases reached the cutoff level in the GJT on ungrammatical DOM items. However, 48 cases in gender OPT and 3 cases in DOM OPT scored above 0.80.

**Appendix D**

***Differences between Feedback Groups on Cognitive Abilities***

Explicit Learning Ability, *F*(2, 103) = 0.83, *p* = 0.436; Implicit Memory Ability, *F*(2, 103) = 0.10, *p* = 0.903; Implicit Learning Ability, *F*(2, 103) = 0.78, *p* = 0.458.

**Appendix E**

***Models Probing the Interactions between Cognitive Aptitudes and Feedback Group***

**The Interaction between Implicit Language Learning Aptitude and Feedback Group**. In the first model, the control group was used as the reference category in dummy coding. In the second model, the explicit feedback group was used as the reference category. Both models produced the same model summary values. These values indicated that the models were significant and explained approximately 10% of the variance, *F*(5, 97) = 2.336, *p* = 0.047, *R*2 = 0.107. The models also produced the same *F* and *p* values for the test of highest order interaction, *F*(2, 97) = 4.152, *p* = 0.019. The test of the highest order interaction was statistically significant, indicating that there was an interaction between Implicit Learning Ability and Group.

*Model 1: Control Group as Reference Group*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *b* | SE | *t* | *p* |
| Constant | 0.101 | 0.042 | 2.380 | 0.019 |
| Implicit vs. Control | 0.088 | 0.060 | 1.462 | 0.147 |
| Explicit vs. Control | 0.026 | 0.061 | 0.420 | 0.675 |
| Implicit Language Learning | 0.029 | 0.043 | 0.671 | 0.504 |
| Implicit Feedback X Implicit Language Learning | 0.092 | 0.060 | 1.522 | 0.131 |
| Control X Implicit Language Learning | 0.082 | 0.061 | -1.342 | 0.183 |

*Model 2: Explicit Feedback Group as Reference Group*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *b* | SE | *t* | *p* |
| Constant | 0.126 | 0.044 | 2.863 | 0.005 |
| Implicit vs. Explicit | 0.062 | 0.061 | 1.013 | 0.314 |
| Control vs. Explicit | -0.026 | 0.061 | -0.420 | 0.675 |
| Implicit Language Learning | -0.053 | 0.042 | -1.227 | 0.223 |
| Implicit Feedback X Implicit Language Learning | 0.173 | 0.060 | 2.879 | 0.005 |
| Control X Implicit Language Learning | 0.082 | 0.061 | 1.342 | 0.183 |

**The Interaction between Explicit Language Learning Aptitude and Feedback Group**. In the first model, the control group was used as the reference category in dummy coding. In the second model, the explicit feedback group was used as the reference category. The model summary values were exactly the same for both models, *F*(5, 97) = 14.044, *p* < 0.001, *R*2 = 0.420. These values indicated that the models were significant and explained 42% of the variance in the data. The values for the test of highest order interaction, which were also the same for both models, *F*(2, 97) = 3.946, *p* = 0.023, showed that there was a significant interaction between Explicit Learning Ability and Group,

*Model 1: Control Group as Reference Group*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *b* | SE | *t* | *p* |
| Constant | 0.044 | 0.025 | 1.732 | 0.087 |
| Implicit vs. Control | 0.172 | 0.036 | 4.804 | 0.000 |
| Explicit vs. Control | 0.293 | 0.037 | 7.860 | 0.000 |
| Explicit Language Learning | -0.002 | 0.026 | -0.091 | 0.928 |
| Implicit Feedback X Explicit Language Learning | -0.052 | 0.037 | -1.415 | 0.160 |
| Control X Explicit Language Learning | 0.054 | 0.037 | 1.444 | 0.152 |

*Model 2: Explicit Feedback as Reference Group*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *b* | SE | *t* | *p* |
| Constant | 0.337 | 0.027 | 12.353 | 0.000 |
| Implicit vs. Control | -0.121 | 0.037 | -3.258 | 0.002 |
| Explicit vs. Control | -0.293 | 0.037 | -7.860 | 0.000 |
| Explicit Language Learning | 0.051 | 0.027 | 1.906 | 0.060 |
| Implicit Feedback X Explicit Language Learning | -0.106 | 0.038 | -2.809 | 0.006 |
| Control X Explicit Language Learning | -0.054 | 0.037 | -1.444 | 0.152 |