Supplementary Materials for: Improved student learning through active retrieval practice and random-sampled exams

KAILI VESIK

DEPARTMENT OF LINGUISTICS, UBC
EMAIL: KAILI.VESIK@UBC.CA

KATHLEEN CURRIE HALL
DEPARTMENT OF LINGUISTICS, UBC
EMAIL: KATHLEEN.HALL@UBC.CA

1. Statistical Models for LING 311 − Accompanies Figure 6

Question: Does the type of LING 200 experience (traditional or randomized) have an effect on the midterm exam, final exam, or course grades in LING 311?

Model 1: lmer(*Grade* ∼ *Type of 200* × *Item* + (1|*Student*))

	Estimate	Std. Error	t-value
Intercept	75.95	2.69	28.23
200-Randomized	-2.03	3.50	-0.58
Item-Midterm	-9.28	2.21	-4.19
Item-Final	-6.96	2.21	-3.15
200-Randomized:Item-Midterm	7.23	2.88	2.52
200-Randomized:Item-Final	2.73	2.88	0.95

Model 2: lmer(*Grade* ∼ *Item* + (1|*Student*))

	Estimate	Std. Error	t-value
Intercept	74.75	1.72	43.55
Item-Midterm	-4.99	1.45	-3.45
Item-Final	-5.35	1.45	-3.70

2. Statistical Model for Course Evaluations − Accompanies Figure 7

Question: Does the type of LING 200 experience (traditional or randomized) have an effect on the student evaluation scores for questions on clarity, effectiveness, fairness, or student interest, in LING 200 itself? (NB: The question on "clarity" in a traditional course served as the reference level.)

Model: lm(*Score ∼ Type of 200 × Prompt*)

	Estimate	Std. Error	t-value	p-value
Intercept	4.71	0.19	39.69	<2e-16
200-Randomized	-0.14	0.17	-0.80	0.42
Prompt-Effective	-0.06	0.17	-0.35	0.73
Prompt-Fair	-0.32	0.17	-1.93	0.05
Prompt-Interest	-0.18	0.17	-1.05	0.29
200-Randomized:Prompt-Effective	0.16	0.24	0.65	0.52
200-Randomized:Prompt-Fair	0.09	0.25	0.36	0.72
200-Randomized:Prompt-Interest	0.09	0.25	0.38	0.71

F-statistic: 1.17 on 7 and 244 DF, p-value = 0.32