

Statistical Analysis on the Best Atmosphere for Memorizing Vocabularies

Chaerin Song

06/04/2021

I. Description

1) Motivation of the Experiment

As an ardent language learner, how to memorize vocabulary as effectively as possible has always been my biggest question. My parents used to tell me to use paper over digital devices and not to listen to music when studying, and to prefer daytime over night for focus tasks. However, I am a night owl, a fond fan of listening to music while studying, and I almost always use an iPad to study these days. In order to test these rumors, and also to find the best focus environment for myself, I conducted a short term experiment on the number of new vocabularies I memorized under different circumstances.

2) Factorial Experiment Design

I am interested in examining three components of memorization efficiency. They are all quantitative factors.

1. *Daytime* (1 for daytime, -1 for nighttime)
2. *Paper* (1 for paper, -1 for iPad)
3. *Music* (1 for presence of music, -1 for silence)

The response variable is the number of new Russian vocabularies that I correctly remember after 20 minutes of concentrated memorization.

The experiment is designed as a 2^k factorial design with $k=3$ and 2 replications. All possible combinations of experimental conditions are:

##	Daytime	Paper	Music
## 1	-1	-1	-1
## 2	1	-1	-1
## 3	-1	1	-1
## 4	1	1	-1
## 5	-1	-1	1
## 6	1	-1	1
## 7	-1	1	1
## 8	1	1	1

Linear Model and Null Hypotheses

Linear model: $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1} x_{i2} + \beta_5 x_{i1} x_{i3} + \beta_6 x_{i2} x_{i3} + \beta_7 x_{i1} x_{i2} x_{i3} + \epsilon_i$
(Note that β_0 is the intercept term.)

x_{i1}	time of the day (day: 1, night: -1)
x_{i2}	writing surface (paper: 1, iPad: -1)
x_{i3}	sound (music: 1, silence: -1)
ϵ_i	error term
y_i	the number of vocabularies correct

As there are 3 main effects and 4 interaction effects in the design, we have 7 null hypotheses.

Null Hypotheses: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$

3) Experiment Procedure

For each run, there are 100 new Russian vocabularies listed with equivalent English translation. With 16 runs of experiment (8 combinations, 2 replications), 16 different lists of 100 Russian nouns are randomly sampled using Excel, from an online Russian-English dictionary csv file (<https://github.com/Badestrand/russian-dictionary>). These lists are either printed out, or saved as pdf documents for iPad viewing.

The number of memorized words is counted through filling out the Russian words on the same list (same order) given the English equivalent. I gave myself 10 minutes for filling out these answers with the consistent paper factor. **In order to remove external factors**, all the lists of vocabularies and the quiz documents are prepared before the whole experiment begins. Also, the music is carefully chosen to be one kind of slow jazz for every run (*jazz link: <https://www.youtube.com/watch?v=uGdYGog1qAM&t=8139s>*).

In order to prevent the order of runs to bias the experiment result, I randomized the order of runs. However, due to the limitation of time given, just for the daytime factor, one run for each daytime was done everyday (daytime at 3 p.m., nighttime at 11 p.m) for 8 consecutive days. For other 2 factors, the order was completely randomized for each day and night runs.

In addition, not reusing the vocabulary list and to recover energy from the previous run was crucial to **guarantee independence in observations**.

Following is the assigned run order, including the replication column to indicate whether the run is for the first or the second replication.

experiment run order				
Order	Rep	Paper	Music	Time
1	Rep 1	1	1	3pm
2	Rep 1	1	1	11pm
3	Rep 2	-1	1	3pm
4	Rep 1	-1	1	11pm
5	Rep 2	1	-1	3pm
6	Rep 2	1	1	11pm
7	Rep 1	1	-1	3pm
8	Rep 2	-1	-1	11pm
9	Rep 2	1	1	3pm
10	Rep 2	1	-1	11pm
11	Rep 1	-1	1	3pm
12	Rep 1	-1	-1	11pm
13	Rep 1	-1	-1	3pm
14	Rep 1	1	-1	11pm
15	Rep 2	-1	-1	3pm
16	Rep 2	-1	1	11pm

Beginning with paper and music on in the afternoon of April 6th, the experiment has followed the randomized order above.

II. Data Analysis

1) Explanatory Data Analysis

Experiment record

Here is the summary table of all 16 runs:

experiment run order				
Run	Daytime	Paper	Music	Correct Answer Count
1	1	-1	1	24.0
2	-1	-1	1	22.0
3	1	1	-1	27.5
4	-1	-1	-1	26.0
5	1	1	1	31.0
6	-1	1	1	28.5
7	1	-1	1	27.0
8	-1	1	-1	29.0
9	1	1	1	27.0
10	-1	1	1	26.5
11	1	-1	-1	26.5
12	-1	-1	-1	25.0
13	1	-1	-1	27.5
14	-1	-1	1	24.0
15	1	1	-1	32.0
16	-1	1	-1	29.0

Summary of the data

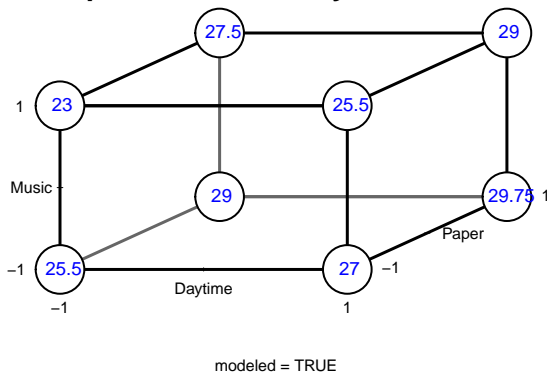
Here is the brief summary of the correct answer counts.

Min	Median	Max	Mean	Variance
22.0	27.0	32.0	27.03	6.68

The mean and the median are almost the same, and the range is quite small.

The cube below shows how this design produces 12 comparisons along the 12 edges of the cube: four measures of the effect of time of day, four measures of the effect of writing surface, and four measures of the effect of presence of music. Note that the value at the each vertex is the average y value of 2 replications.

Cube plot for vocabulary memorization



2) Factorial Effects

Main effects

Main effect of a factor measures the average differences of the factor's two levels.

Factor	Estimate	P-value
Daytime	0.78	0.1311
Paper	1.78	0.0050
Music	-0.78	0.1311

Interaction effects

Interaction between two factors is the average difference between each factor's average difference. The difference between these two interactions measures how consistent the Daytime:Paper interaction for two levels of Music. Half this difference is defined as the three factor interaction of Daytime, Paper, and Music.

Factor	Estimate	P-value
Daytime:Paper	-0.22	0.6503
Daytime:Music	0.22	0.6503
Paper:Music	0.22	0.6503
Daytime:Paper:Music	-0.03	0.9480

Estimated Variance of Each Conditions

As I conducted an experiment with 2 replications, I can use the difference of the two measures to get a variance of each conditions.

Daytime	Paper	Music	s_i^2
1	1	1	8
1	1	-1	10.125
1	-1	1	4.5
1	-1	-1	0.5
-1	1	1	2
-1	1	-1	0
-1	-1	1	2
-1	-1	-1	0.5

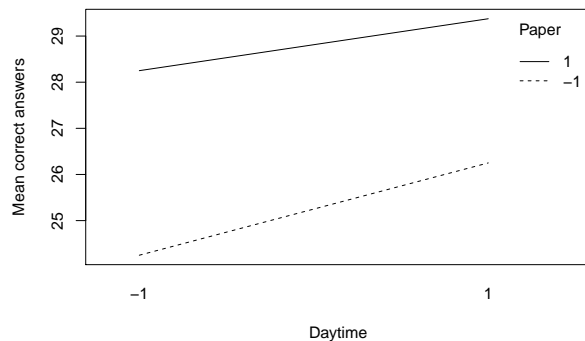
Some of these variances are larger than others, but it is unnecessary to put a lot of meaning into each variance, as there were only 2 replications.

3) Interpretation

Interaction Plots

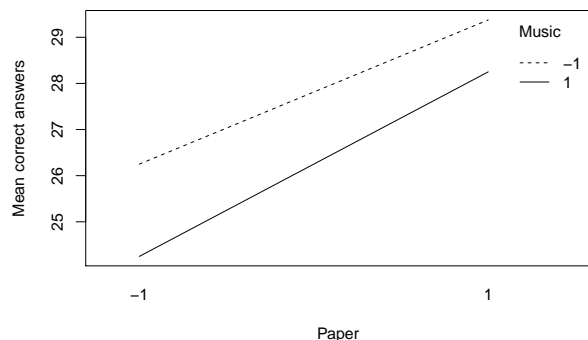
For interpretation, here are interaction plots where they show the mean counts of correct vocabularies for each pair of factors.

1. Daytime:Paper Interaction plot



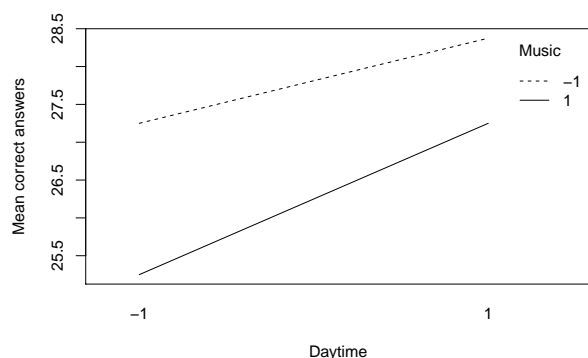
There is no crossing, and two lines are close to parallel. This indicates that there is little sign of interaction between Daytime and Music factors.

2. Paper:Music Interaction plot



Similarly for Paper:Music interaction, two lines are almost parallel, and there is no crossing happening. This indicates little interaction between Paper and Music factors.

3. Daytime:Music Interaction plot



Compared to the past two interaction plots, two lines in this Daytime:Music interaction plot is less parallel. Still, there is no crossing. This indicates that although there could be a minor interaction of Daytime and Music factors, it is not significant.

In summary, it is hard to say that there is any non-trivial interaction in our results. This result is consistent with our high p-values from the interaction effects as well. As there is no evidence that factors interact with each other, we can interpret the main effects individually.

Linear Model

The estimated coefficients are $\frac{1}{2}$ of the factorial estimates, and the intercept β_0 is the sample mean. Based on this, here are our estimated coefficients for all the variables, including the intercept term with their p-values.

Variable	Estimate	P-value
β_0	27.03	0.0000
β_1	0.3906	0.1311
β_2	0.8906	0.0050
β_3	-0.3906	0.1311
β_4	-0.1094	0.6503
β_5	0.1094	0.6503
β_6	0.1094	0.6503
β_7	-0.0156	0.9480

Thus, our linear model is:

$$\hat{y}_i = 27.03 + 0.3906x_{i1} + 0.8906x_{i2} - 0.3906x_{i3} - 0.1094x_{i1}x_{i2} + 0.1094x_{i1}x_{i3} + 0.1094x_{i2}x_{i3} - 0.0156x_{i1}x_{i2}x_{i3} + \hat{\epsilon}_i$$

Note that with significance level of 0.05, only the intercept term and the coefficient of the Paper factor are significant. All 4 interaction variables' coefficients, especially, have very high p-values, which shows that there is no significant interaction between the factors.

The only significant factor is the Paper factor. If everything remains the same but the writing surface, then the number of vocabularies memorized are estimated to be different by 0.8906.

Going back to the null hypotheses from the beginning, based on the results above, we reject the null hypothesis at 0.05 significance level that there is no difference in the number of vocabularies correctly memorized between studying on paper versus studying on iPad. For every other null hypotheses, we fail to reject at 0.05 significance level.

Confidence Interval

Here I will calculate 95% confidence intervals of each factorial effect. Below is the table of confidence intervals for each effect.

Effect	Lower	Upper
Daytime	-0.580	3.705
Paper	1.420	5.705
Music	-3.705	0.580
Daytime:Paper	-2.580	1.705
Paper:Music	-1.705	2.580
Paper:Music	-1.705	2.580
Daytime:Paper:Music	-2.205	2.080

Each effect has the confidence interval with length of about 4, and other than the Paper factor, all confidence intervals include 0, which means that we are not sure more than 95% confidence that any effect other than the Paper factor would be either positive or negative.

III. Conclusions

Our linear model indicates that on average, the number of vocabularies memorized will be increased by 0.3906 if done in daytime over nighttime, 0.8906 when studying on paper over iPad, and 0.3906 when studying in silence over jazz. Interaction effects are inferred to not only be small, but also not significant at all (p-values are over 0.6).

Thus, we can conclude that it is more efficient to memorize vocabulary when I do it during the daytime over nighttime, when I study on paper over iPad, when in silence over jazz music. Hence, the optimal experimental condition is to study during daytime using paper with no music playing in the background. With this conclusion, I learned how to make the best atmosphere for memorizing Russian vocabularies.

IV. Discussion and Limitations

Future Improvement

One of the reason why the p-values of many effects were not small enough to be significant with 95% confidence level was the number of observations. For each set of condition out of total 8, there were only 2 measurements due to the limitation of time. In the future, a similar experiment could be done with more replications and possibly more experimental units (currently 1) and factors. While I was doing each run, I realized that there could be more important factors in the efficiency of memorizing vocabularies, such as the hour of sleep the night before, the hour since the last meal, or even the brightness of the room. These can be further introduced in later when I get a chance.

Source of Bias

As mentioned in the previous part, there are many other factors that I consider to be possibly important. Because of personal matters, a lot of these conditions were not well controlled, which could have been the source of bias. Also, sometimes it was hard to start each run exactly at 3pm and at 11pm. This inconsistency in time could also have affected the result, especially on the Daytime factor. My daily condition could have been a big factor as well, but this was quite tricky to maintain consistently.

Limitations of the Study

Considering that the experiment was done under very limited time, limited experimental unit, and limited amount of factors, there are a lot of limitations of the study. First, the only experimental unit was just myself. It is hard to extrapolate the result of the experiment to general students. Second, it was done only with Russian vocabularies, specifically nouns. Thus, we are not sure if the result would be consistent on other languages or even other grammatical categories in Russian language. Lastly, we can only apply the result of Music factor on a slow jazz with no lyrics. If I played other types of music, the result might have been different.