

# **DETERMINING THE SIGNIFICANT FACTORS THAT AFFECT THE TYPING SPEED**

Arvind Ramkumar, Sajil Odanath, Viren Bhanushali

## **EXECUTIVE SUMMARY**

Typing is a necessity in the creation of any worded document. A typist should have a good speed and a low error record. These are the factors that are considered when a person is hired for a typing job. Typing is used in various fields from typing books to code programs. But to have a good speed there are factors that will affect the typing speed.

The experiment that we conducted is one such experiment, where the factors like different languages, types of keyboards and some sort of distraction were part of the experiment. There are other factors that influence the typing, but we consider these to be the main factors. The experiment was performed to maximize the typing speed. Two subjects were chosen and made to write a paragraph of different languages. This factor was also combined with the other factors mentioned above and a total of eight combinations were run for each subject.

The results obtained from this experiment shows that only two out of the three factors are significant factors that contribute to final model. Also, the interaction between these two factors is a significant factor for the final model.

The final model includes the terms type of passage, the distraction and the interaction between these two terms. The results obtained from the experiment are included in this report. These kinds of experiments are helpful for managers when hiring new employees according to the kind of typing job.

## **Problem Definition**

Typing speed is considered as one of the important skills for any individual in the company. Now considering an engineering student, touch typing helps in coding, writing reports, emailing etc. Typing speed increases the efficiency of working as one can concentrate more on the context rather looking at the keyboard.

The ultimate aim of the experiment is to determine the factors that would maximize the typing speed of an individual on a computer. Typing speed depends on a number of factors and thus, this experiment helps us to identify the factors that majorly affect any individual's typing speed.

Response Variable (Words per min)	Operating Range (Words per min)	Precision	Objective
Typing speed	25-60	Least count of 1 word	Maximization

## **Selection of Response Variable**

Typing speed (words per min) is selected as the response variable for the experiment. It is measured by making individuals type a specific paragraph and recording the number of words completed in a minute. Number of words were visible in the Microsoft word software and we ensured that spelling mistakes were not considered as a word.

## **Selection of Factors**

### **Paragraphs in different languages**

For the experiment we have decided to test the subjects providing them a paragraph to write in different languages. This will help test the typing speed as the individual may not be proficient in the language being tested.

### **Distraction during the experiment**

We would be testing the subject in which they will performing under some sort of distraction (music in the background) to see if this factor affects the typing

### Keyboard type

Typing speeds also depend on the type of keyboard being used. There are soft and hard keyboards which can affect the speed of typing.

Factors	Type	Levels (2 Levels)	
Different Paragraph	Categorical	-1 = English	1 = French
Distraction	Categorical	-1 = No Music	1 = Music
Keyboard type	Categorical	-1 = Soft	1 = Hard

## **EXPERIMENTAL SETUP**

### Choice of Experimental Design

With the above design parameters, we propose conducting a  $2^3$  randomized complete block design. We would be using 2 replicates in the design and blocking would be done as per the replicates. We propose to Block the Number of Individuals as variability that might affect the result. The choice of blocking is also attributed to eliminating the known and controllable factor that is Individual Capability in the particular experiment. Thus, we can systematically eliminate its effect on the statistical comparisons among treatments. Hence, Three Factors, each with Two levels and two replicates are chosen for the design.

The choice of number of replicates had been decided with the help of JMP Software.

A replicate size of two indicates a 2-standard deviation of 93.7% and replicate size of 3 indicates 99.6 % at 95% confidence interval. Since a higher standard deviation reflects Design of Engineering Experiments better difference in variability, we therefore have performed a  $2^3$  Randomized complete block design.

### Designing the Experiment

There are various factors that affect the randomization assumption in the experiment. Several measures were taken to minimize the problems that will be faced while conducting the experiment. To avoid the person from getting used to the paragraph they type, they were not given a same passage to type over and over again. Since it is just the Words per minute count, this was not a problem. Since the randomization is within the blocks, the individual did select their experiment

combination at random. In order to avoid the fatigue of the individual, they were given periodical rest between each of the runs.

For convenience, to insert the values in the JMP table, we sorted the matrix from left to right.

	Pattern	Block	Type Of Paragraph	Distraction	Type of Keyboard	Typing Speed (WPM)
1	----	1	-1	-1	-1	•
2	----	2	-1	-1	-1	•
3	+-	1	1	-1	-1	•
4	+-	2	1	-1	-1	•
5	++	1	-1	1	-1	•
6	++	2	-1	1	-1	•
7	++	1	1	1	-1	•
8	++	2	1	1	-1	•
9	++	1	-1	-1	1	•
10	++	2	-1	-1	1	•
11	++	1	1	-1	1	•
12	++	2	1	-1	1	•
13	++	1	-1	1	1	•
14	++	2	-1	1	1	•
15	++	1	1	1	1	•
16	++	2	1	1	1	•

Factors	Levels
Different Paragraph	-1 = English , 1 = French
Distraction	-1 = No Music , 1 = Music
Keyboard type	-1 = Soft, 1 = Hard

## **STATISTICAL ANALYSIS OF DATA**

### Model Analysis

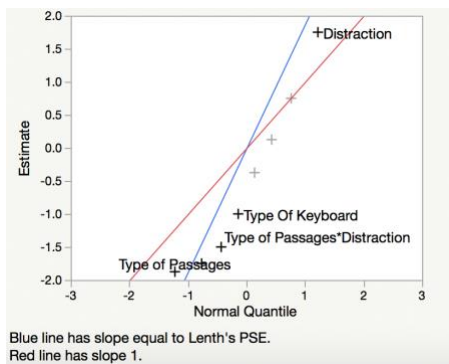
Initially, with all the three factors (Type of Paragraph, Distraction and Type of Keyboard), we fit the full model with all the two factor and three factor interaction terms.

The effect Summary and the parameter estimates of the full model is given below.

Effect Summary			
Source	LogWorth		PValue
Type of Passages	1.833		0.01467
Distraction	1.703		0.01982
Block	1.703		0.01982
Type of Passages*Distraction	1.435		0.03673
Type Of Keyboard	0.887		0.12972
Type of Passages*Type Of Keyboard	0.522		0.23682
Type of Passages*Distraction*Type Of Keyboard	0.267		0.54022
Distraction*Type Of Keyboard	0.078		0.83620

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	32.625	0.582482	56.01	<.0001*
Type of Passages	-1.875	0.582482	-3.22	0.0147*
Distraction	1.75	0.582482	3.00	0.0198*
Type Of Keyboard	-1	0.582482	-1.72	0.1297
Block[1]	-1.75	0.582482	-3.00	0.0198*
Type of Passages*Distraction	-1.5	0.582482	-2.58	0.0367*
Type of Passages*Type Of Keyboard	0.75	0.582482	1.29	0.2388
Distraction*Type Of Keyboard	0.125	0.582482	0.21	0.8362
Type of Passages*Distraction*Type Of Keyboard	-0.375	0.582482	-0.64	0.5402



From the above table of Effect Summary and Parameter Estimates and the Normal Plot, it is quite evident that the three-factor interaction term is not significant. And also, for the 95% confidence, there are so terms that are not significant. The only significant terms are the Type of Passage (A), Distraction (B), and the interaction between these two (A\*B). The Type of keyboard does not have any significant impact on the response (WPM).

So, Now, we fit a reduced model with all the significant terms to obtain the prediction equation.

#### Effect Summary

Source	LogWorth	PValue
Type of Passages	1.975	0.01059
Distraction	1.816	0.01528
Block	1.816	0.01528
Type of Passages*Distraction	1.498	0.03175

#### Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	32.625	0.610048	53.48	<.0001*
Type of Passages	-1.875	0.610048	-3.07	0.0106*
Distraction	1.75	0.610048	2.87	0.0153*
Block[1]	-1.75	0.610048	-2.87	0.0153*
Type of Passages*Distraction	-1.5	0.610048	-2.46	0.0317*

Furthermore, to verify the model we make use of the ANOVA Estimation.

#### Summary of Fit

RSquare	0.743891
RSquare Adj	0.65076
Root Mean Square Error	2.440194
Mean of Response	32.625
Observations (or Sum Wgts)	16

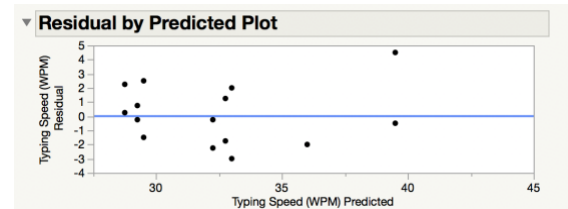
#### Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	190.25000	47.5625	7.9876
Error	11	65.50000	5.9545	Prob > F
C. Total	15	255.75000		0.0028*

The Analysis of Variance has a high F ratio and the p value is also very low. So, this model is proved to be significant.

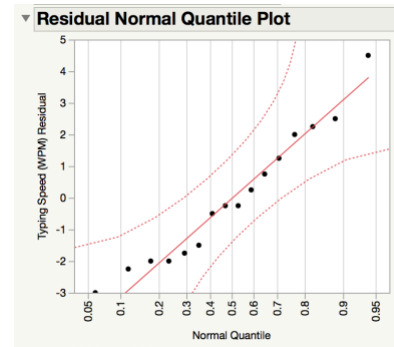
## Residual Analysis

### Residual by Predicted Plot



There is no specific pattern found in the Residual By Predicted plot. This plot helps us in understanding the Assumption of the Constant Variance. Since there is no specific pattern found, we could conclude that constant variance assumption is not violated.

### Normality Quantile Plot



The Residual Normal Quantile Plot helps in validating the Normality Assumption. Here, all the observations pretty much seem to lie along the fitted line. So, it could be concluded that the normality assumption is valid.

## Transformation

Although Residual Plot looks perfectly fine, the RMSE of the model is high (2.440194), and also it is possible to perform few transformations on the response, to make the model even more better. We performed few of the transformation including the Square root, Log and the reciprocal. Out of all the Transformations that was used, the log transformation did yield the better result. The Summary of Fit and ANOVA is given below.

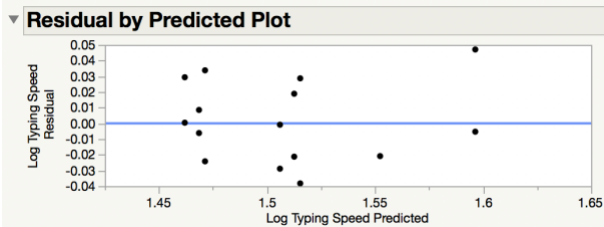
Summary of Fit	
RSquare	0.752523
RSquare Adj	0.662531
Root Mean Square Error	0.029768
Mean of Response	1.510585
Observations (or Sum Wgts)	16

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	0.02963897	0.007410	8.3621
Error	11	0.00974717	0.000886	Prob > F
C. Total	15	0.03938614		0.0024*

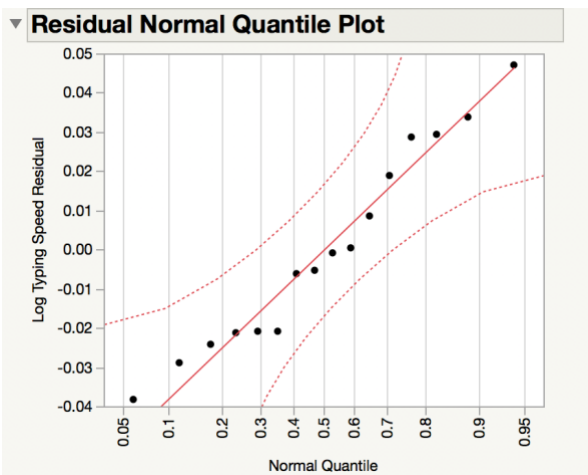
The RMSE of the Log transformed model is very much less when compared to untransformed model (0.029768) and the residual plot also seems very much better than the original model and not violating the Assumptions also.

Residual Plot of the transformed

Residual By Predicted



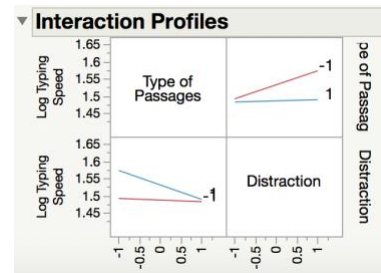
Normal Quantile Plot



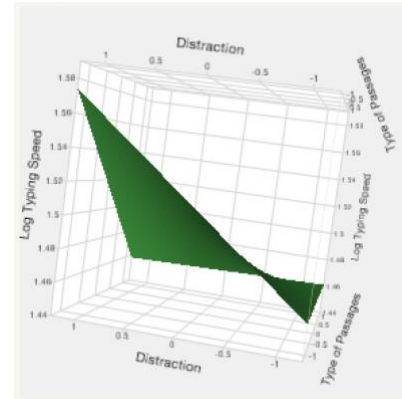
Thus, to know the prediction we look at the Parameter estimates.

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	1.5105852	0.007442	202.98	<.0001*
Type of Passages	-0.023286	0.007442	-3.13	0.0096*
Distraction	0.0218942	0.007442	2.94	0.0134*
Block[1]	-0.022017	0.007442	-2.96	0.0130*
Type of Passages*Distraction	-0.018603	0.007442	-2.50	0.0295*

To know more about the Interaction term, we utilize the Interaction plot.



The significant interaction between the Type of Passage and the distraction will help us to understand the combined effect of the two. So, it is mandatory for us to consider about this effect also, and the response surface also changes accordingly.



The Curvature in the Surface Plot is due to the interaction between the factors – Type of Passage and Distraction.

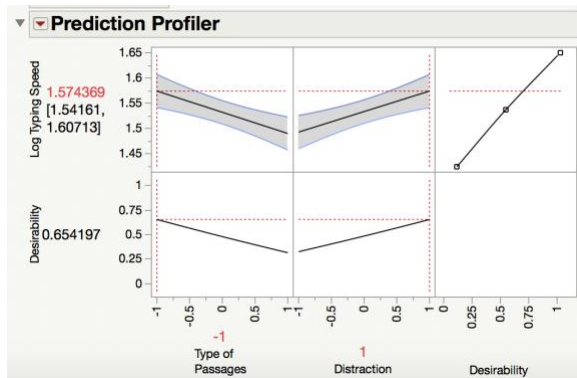
Prediction equation

$$y = 1.5105 - 0.02386*A + 0.021894*B - 0.01860A*B$$

### Maximum Desirability Conditions

In Order to maximize the Typing speed, we analyze the prediction profiler plot to find the values in Interaction term, typing speed is higher when the type of passage is English, and Distraction is high level.

Factor	Level	Value of the Level
Type of Passage	-1	English
Distraction	1	Music



### CONCLUSION

The objective of our experiment was to identify which factors significantly affect the Typing speed of an individual. The Experiment was conducted in order to train the employees in the industry to optimize their experience with efficient typing. Based on the analysis, we identified that Type of Passage (A), Distraction (B) and the interaction between these two (A\*B) were significant. This experimental data was run through analysis and confirmed by Normal Probability plot, ANOVA and Regression. Since we have interaction term, so looking at the interaction profile, we concluded that the result was better while typing the English passage (Low Level) along with Distraction (MUSIC ON).

### RECOMMENDATION

The Experiment we conducted gives an insight about the significance of the factors, but it is important to figure out the optimum values of these factors in order to maximize the Typing speed. Apart from these factors, we can also look at the ergonomics of the keyboard and the typing pattern of an individual, but this was out of scope. We can perform response surface modelling to find the optimal values. Once we gain the right knowledge of biomechanics, we can improve the results of this experiment. Also, we have considered only 2 levels of passages, but the results might differ when considering languages with difficult words or Mathematical expressions. So, in future we will consider all the remaining and improve our range of the experiment using the learning from the current design.