Differences in Successive Measurements of Reaction Time throughout the Day

Vivian Ngo

December 3, 2018

Abstract

Reaction time can play a large role in determining when traffic incidents are more likely occur, how much time it would take for civilians to evacuate a building that is on fire, how frequently a dodegball player can dodge a ball, and more. This paper discusses how average reaction time changes throughout the day, specifically for the 2018-2019 cohort of STA490 students. Assuming four spaced out measurements throughout the waking hours of an individual, on average, the first measurement of reaction time is significantly higher than that of the second measurement and the third measurement, and the second measurement of reaction time is significantly lower than that of the fourth measurement. In addition, this difference does not change for individuals who are morning or evening persons, or those who are neither.

Introduction

The objective of this project is to investigate the diurnal pattern of reaction time, that is, the pattern of reaction time throughout the waking hours for an individual. In particular, the focus is to determine if an association exists between the measurement of the day and reaction time. Comparisons are made between the first, second, third, and fourth measurements of reaction time, all equally spaced out at approximately four to six hour intervals. The data used in this study was collected from 39 students from the 2018-2019 cohort of students from STA490 at the University of Toronto.

This investigation was performed because reaction time can play a large role in critical and non-critical situations. Reaction time is important when driving, cooking, playing sports, performing surgery, writing a test, and in many other activities as well. Knowing about patterns in reaction time can play a large role in determining when it is safer to perform different activities. The rest of this report will describe the data recorded, procedures used to collect the data, methods used to analyze the data, and the results of the study.

Data Summary

Protocol

40 students from the 2018-2019 cohort of STA490 students were asked to perform the reaction time test from https://questionnaire.censusatschool.ca/practice/reaction four times each day for two days. The two days were to differ by busyness, as determined by each individual student, in order to get data that would be less skewed by specific schedules on certain days. As well, the tests were to be taken at four to six hour intervals, with the first test taken within an hour of waking up and the last test taken within an hour of going to sleep. Subjects were not required to take the tests at specific predetermined times because it was assumed that different students are awake during different times of the day and have different sleeping patterns. Because the focus of this study is the reaction time pattern for each individual's waking hours rather than the reaction time pattern in a 24-hour day, subjects were asked to perform the tests according to their personal waking hours.

For each measurement, subjects were only allowed to perform the test once without practice runs and were required to record the value from that one test. Repeated tests for each measurement were not allowed in this study, to prevent the possibility of learning effects. As well, for the measurements to be consistent, each subject was told to use their dominant hand and the same electronic device to perform the tests for all eight measurements. Subjects were also asked to wear any visual aids that they required to see clearly, and to be in a comfortable environment when performing the test.

Table 1: Variable Descriptions

Variable	Description
ID	The unique ID number for every student, ranging from 1 to 40
Measurement Number	Number of the measurement. Ranges from 1 to 8
Raw Time	The 24 hour format time when the data was recorded
Day Type	1 indicates busy day, 0 indicates light day
Stimulant	1 indicates a stimulant was taken, 0 indicates otherwise
Fatigue	Fatigue measured on the Samn-Perelli 7-pt scale. Ranges from 1 to 7
Hunger	Hunger measured on the hunger scale, value decreasing with hunger level
Illness	1 indicates illness, 0 indicates otherwise
Sleep	The hours of sleep that was had the night before the observation
MEQ	The Morningness Eveningness Questionnaire score, ranging from 16 to 86
Reaction Time	The reaction time that was measured (in seconds)
Break Protocol	1 indicates protocol was broken, 0 indicates otherwise
MEQ Category	Classfication of MEQ score into Evening, Morning, or Neither.
Measurement of the Day	If the observation was first, second, third, or fourth observation of the day

Table 2: Categorical Variables

Day Type	#	Stimulant	#	Fatigue	#	Hunger	#	Illness	#	MEQ Category	#
0	152	0	247	1	20	1	7	0	259	Evening	112
1	155	1	41	2	43	2	2	1	32	Morning	8
NA	13	NA	32	3	84	3	42	NA	29	Neither	176
				4	61	4	83			NA	24
				5	57	5	76				
				6	23	6	54				
				7	6	7	22				
				NA	26	8	4				
						9	4				
						NA	26				

Table 3: Continuous Variables

	Sleep	MEQ	Reaction Time	Time
Minimum	4.0	32.0	0.10	0.00
Mean	7.9	44.4	0.40	14.00
Maximum	12.0	60.0	0.85	23.93
NA's	44.0	24.0	18.00	33.00

Data Cleaning

Before any analysis was performed, steps were taken to make the data consistent and clean. In particular, if an MEQ was reported as one of the categorical names instead of a number then it was replaced by the average number in that category. Another variable was also created using MEQ, called "MEQ Category" and it is a categorical variable representing MEQ, as opposed to the numerical value that was reported. Any MEQ that was in the range of 16-41, 42-58, and 59-86 was represented as "Evening", "Neither", or "Morning", respectively.

Also, Illness was initially a variable that subjects could write anything for. For example, subjects could write "cough", "flu", or anything else they wanted. Because this data structure would be very messy to deal with, Illness was changed to a categorical variable with two levels, namely 1 or 0, indicating that an illness was present or not. In the transformation of this variable, any illness that was reported, regardless of severity, was recorded as an illness. This decision was made because severity of an illness is subjective so it was assumed that any illness recorded by the subject represented something uncomfortable or unusual.

Methods (for the Analysis)

Each individual was asked to perform the reaction time tests on a busy and a light day, and those measurements were treated as replicates, as opposed to variables of interest in this study. So, the response variable that was analyzed was the mean of these replicates.

As per the protocol, each subject was expected to have one mean reaction time for each of the four measurements throughout the day. Thus, each subject would have four observations. If a subject did not have any values recorded for the same measurement time on both days, it was not possible to calculate a mean reaction time for that measurement number. These individuals were excluded from the study as they did not produce four observations for analysis. For the subjects that were kept in the analysis, each subject had repeated measurements. Thus, it was reasonable to model reaction time against the measurement of the day using a linear mixed-effects model, with subject ID as the random effect.

An assumption was also made that fatigue was an adequate summary for variations in sleep, hunger, and illness since level of sleep, hunger, and illness typically affect fatigue levels. After this assumption, the variables of interest that remained were MEQ, Fatigue, Measurement of the Day, and reaction time. The first model built was a linear mixed-effects model using Measurement of the Day, MEQ Category, and Fatigue as fixed effects and Subject ID as a random effect.

After the linear mixed-effects model, the nonparametric Wilcoxon signed-rank test and Bonferroni correction were used to detect significant differences between the average reaction times for pairs of different measurements throughout the day. The Wilcoxon signed-rank test was a suitable test to use because the reaction time measurements within each level of measurement numbers were paired, differences were symmetric, and the response variable is continuous. Most importantly, this test does not make assumptions on the underlying model of the data, whereas the linear mixed-effects model does. Using the two models would allow for a better understanding of the data.

The linear mixed-effects model and the Wilcoxon signed-rank test did indeed produce results that were consistent with each other.

Results

The first model that was built assuming reaction time as a function of Measurement of the Day, MEQ Category, Fatigue, and Subject ID, produced results that were consistent with the hypothesis that reaction time is associated with Measurement of the Day. In fact, the ANOVA table showed that both Measurement of the Day and Fatigue were associated with reaction time (with p-values of 0.0009 and <0.0001, respectively).

On average, the first measurement of the day has reaction time that is 0.022, 0.028, and 0.033 seconds greater than the second, third, and fourth, with fatigue and MEQ levels held constant (with p-values 0.03, 0.01, and 0.00). Reaction time is also 0.020 seconds greater, on average, when an individual is fatigued, compared to when they are not fatigued, holding other variables constant (with corresponding p-value of 0.00).

Furthermore, the Wilcoxon signed-rank test paired with Bonferroni correction showed that there is significant difference between the average reaction times for the first and second measurement, first and third measurement, and second and fourth measurement of the day (with uncorrected p-values of 0.001, 0.007, and 0.008). In addition, this pattern is not significantly affected by whether an individual is a morning person, evening person, or neither (p-value 0.6216).

Differences between the mean reaction times of the four Measurement of the Day levels can be seen in Figure 1 below. In particular, there is little overlap between the interquartile ranges (boxes) of reaction time for Measurement 1 and Measurement 2 of the day. This also applies to Measurement 1 and Measurement 3. The lack of overlap is consistent with the findings that, on average, the reaction time of the first measurement of the day is different from the reaction time of the second and the third measurements.

Figure 2 also visualizes the changes in reaction time throughout the four measurements of the day. The smoothing curve in Figure 2 is concave upward, suggesting that reaction times are higher at the first and last measurements. This certainly is consistent with the results of the analysis as well.

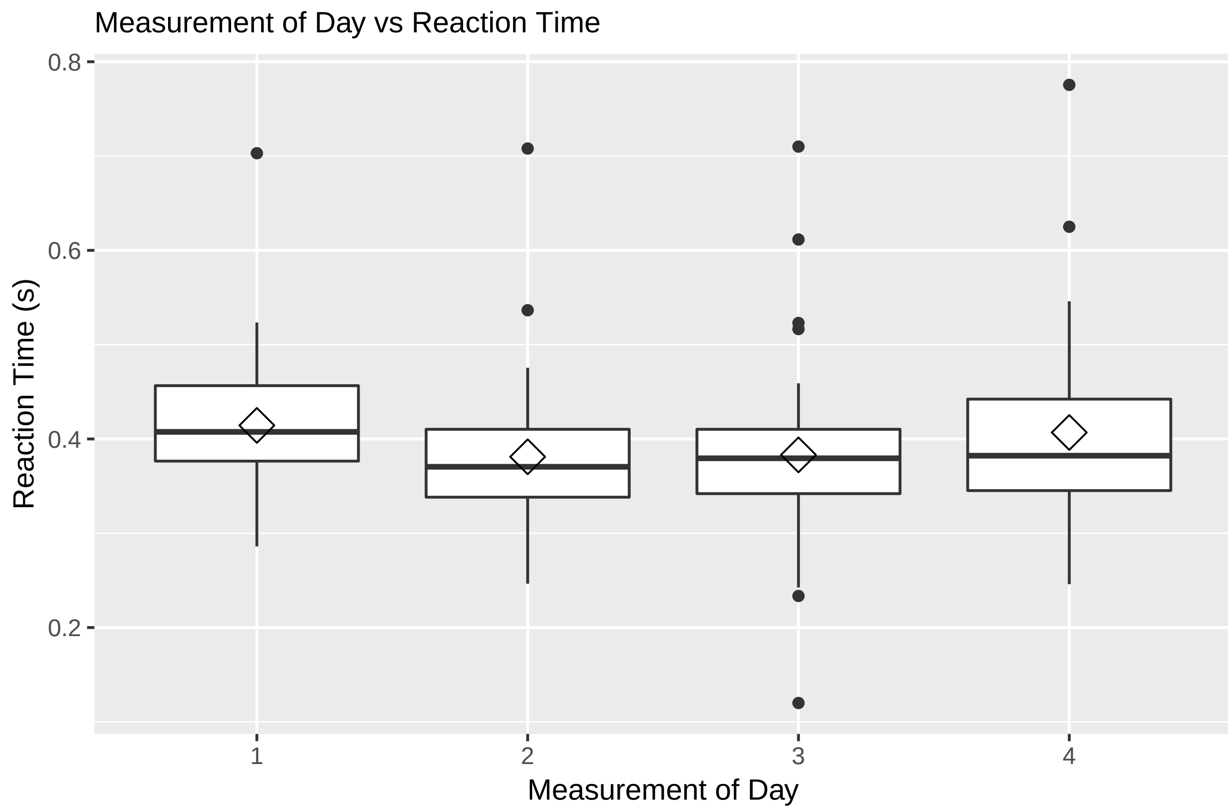

Figure 1

In this experiment, the mean reaction times for different measurements are listed in Table 4 below:

Table 4: Mean Reaction Time at Different Measurements of the Day

Measurement of the Day	Mean Reaction Time (s)
1	0.414
2	0.381
3	0.383
4	0.407

The linear mixed-effects model indicated that there is a significant difference between at least one pair of measurements and the Wilcoxon test further suggests which pairs are statistically different. From the tests and the means found in Table 4, on average, the reaction time of the first measurement of the day is higher than the reaction time of the second and the third measurements. In addition, the second measurement of reaction time in a day is on average lower than the reaction time recorded at the fourth measurement of the day.

Conclusion and Discussion

All in all, there exists a statistically significant difference in average reaction times depending on whether the reaction time is the first, second, third, or fourth measurement taken in a day. These results reveal a slightly concave upward pattern in diurnal reaction time and can help decision makers to determine optimal times for different activities.

However, there are limitations to this study such as the possibility of confounding variables. For example, the time span of this study could be a confounding variable. The data for this study was collected over just one week in 2018 and it is unclear if that week could have affected the diurnal pattern of reaction time. Since our interest is diurnal pattern in general, the data that was collected may not be enough to answer the present research question. Perhaps it would be better in the future to collect observations over a larger time span.

In addition, because the subjects of the study knew that they would be analyzing the data, there could be bias in the data. For example, subjects could have induced a trend in the data or recorded false data to be consistent with a predetermined hypothesis. A possible way to reduce the risk of this issue is to not let students know that they will be analyzing the data or to have students collect the data without knowing the purpose.

Additional studies can also be done to further explore diurnal patterns of reaction time. First of all, the sample that this study utilized only consists of 39 individuals, all of which are Statistics students at the University of Toronto. To generalize the results of this study, further studies should be done with a larger sample of individuals. It would also be interesting to analyze subjects that sleep at different times throughout the 24 hour day. Because this study's subjects were university students, they were subject to class schedules determined by the University of Toronto. If we could also study subjects that work night shifts, or are not in school, this would widen the scope and generality of the results.

In summary, the analysis for this study suggests that reaction time differs throughout the day. Assuming four spaced out measurements throughout the waking hours of an individual, on average, the first measurement of reaction time is significantly higher than that of the second measurement and the third measurement, and the second measurement of reaction time is significantly lower than that of the fourth measurement. This result can be useful when determining optimal times for different activities that improve performance, reduce risk, and more. Future studies on larger samples is suggested to further investigate the diurnal pattern of reaction time for larger populations.