camelCase vs. kebab-case

Experiment 2, Experimentation & Evaluation 2020

Abstract

I want to conduct an experiment which aims to determine which style of naming is better: separating words with a dash character (kebab-case) or using capital letters to denote the start of each word (camelCase). 5 participants tried to identify the correct identifier in both naming systems and with different number of distractors and different number of words per identifier. I analysed the results and the limits of this experiment that indicates the kebab-case style as best suited for readability and accuracy.

1. Introduction

This experiment aims to determine which style of naming is better: separating words with a dash character (kebab-case) or using capital letters to denote the start of each word (camelCase). I want to perform this experiment because it can potentially improve the readability of source codes, which translates into faster programming and faster understandability of programmers that have to work on the code after it has been written. In the past there has been some research in *natural language reading* that found that spaces help reading and removing spaces can slow down reading by 10-20% [Epelboim et al. 1997]. This experiment will try to confirm this result by giving an answer based on the measurement of the speed and accuracy taken from the tests performed by the participants.

Hypotheses:

Separating words with a dash character is more similar to our way to read. The general convention for writing is using spaces between words, and a dash character introduces white space between the words. For this reason I expect the kebab-case to produce faster and more accurate results than the camelCase when increasing the number of identifiers shown to the participants and the length of the phrase.

2. Method

2.1 Variables

Independent variable	Levels
Naming style	camelCase kebab-Case
Number of identifiers	4 8
Number of words in the identifier	2 4

Dependent variable	Measurement Scale
Time	Parametric Ratio scale
Accuracy (Right or Wrong)	Non-parametric Nominal Scale

Control variable	Fixed Value
Screen dimension	16 inch
Computer position	Same position on the desk for every participant

Blocking variable	Levels
English level	Advanced Intermediate Beginner

2.2 Design

Type of Study (check one):

☐ **Observational Study**	☐ **Quasi-Experiment**	☑ **Experiment**

Number of Factors (check one):

☐ **Single-Factor Design**	☑ **Multi-Factor Design**	☐ Other

Between vs. Within (check one):

☐ **Between Group Design** (independent measures)	☑ **Within Subject Design** (repeated measures)	☐ Other

This experiment is a Multi-Factor Design with the independent variables being the naming style, the number of identifiers and the number of words in the identifier. Due to the fact that I am not testing every size of identifiers and and with every number of identifiers shown to the participant, this is a Fractional factorial experimental design. I have chosen this method, which takes into consideration different numbers of identifiers shown, in order to challenge the participants a little more, so that if there is a difference between the two styles, this can come up with a more difficult task. Trying this experiment not only with 2 words per identifier but also 4 makes sure to cover more than 90% of the number of words per identifier in a source code. The ability to read English should not produce any difference between the two different styles, it could only impact the overall time for both kebab-case and camelCase.

For each participant:

Time measurement starts	User clicks on word camelCase 4 identifiers 2 words	Time measurement ends	Right or Wrong
Time measurement starts	User clicks on word camelCase 8 identifiers 2 words	Time measurement ends	Right or Wrong
Time measurement starts	User clicks on word camelCase 4 identifiers 4 words	Time measurement ends	Right or Wrong
Time measurement starts	User clicks on word camelCase 8 identifiers 4 words	Time measurement ends	Right or Wrong
Time measurement starts	User clicks on word kebab-Case 4 identifiers 2 words	Time measurement ends	Right or Wrong
Time measurement starts	User clicks on word kebab-Case 8 identifiers 2 words	Time measurement ends	Right or Wrong
Time measurement starts	User clicks on word kebab-Case 4 identifiers 4 words	Time measurement ends	Right or Wrong
Time measurement starts	User clicks on word kebab-Case 8 identifiers 4 words	Time measurement ends	Right or Wrong

2.3 Participants

ID	Gender	Age	English Level	Other
1	Male	25	Advanced	Friend
2	Female	25	Intermediate	Friend
3	Male	25	Advanced	Friend
4	Male	21	Intermediate	Family member
5	Female	10	Beginner	Acquaintance

2.4 Apparatus and Materials

Experiments carried out at home using the same set up for every participant:

- Same comfortable position at the desk
- Same position of the computer on the desk (roughly same distance from the eyes of every participant)
- Same instructions given to everyone
- Same computer used:

- Same program used to carry out the test: self made Python (3.8.2 64-bit) program written using Jupyter 6.0.3. The time measurements are taken with the Time module when showing the words and when clicking on a word. The time is then printed in the console..

2.5 Procedure

I invite the participants to sit comfortably at my desk, computer sitting in front of them at the same distance for everyone (15 cm from the edge of the desk). I then start the program explaining what will happen, the participant has to complete a practise task in order to understand how it works. After this the real experiment starts, first is tested the camelCase style 4 times:
- 4 identifiers and 2 words,
- 8 identifiers and 2 words,
- 4 identifiers and 4 words,
- 8 identifiers and 4 words.

The times and results are recorded. The participant then goes through the same tests but with the kebab-case style. The times and results are recorded.

3. Results

3.1 Visual Overview

Here I present the results of the times of the experiment (in seconds and precision goes up to 4 digits after the comma):
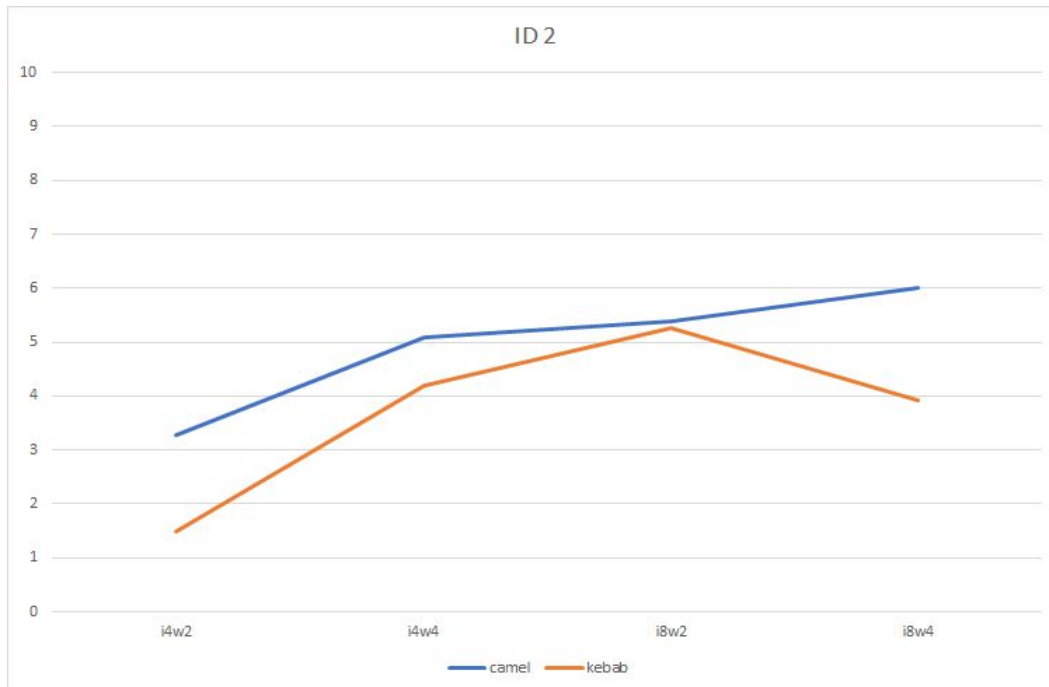
Style	ID	4 identifiers 2 words	4 identifiers 4 words	8 identifiers 2 words	8 identifiers 4 words
camelCase	1	1.6756	4.0604	2.3853	2.7815
	2	3.2754	5.0966	5.3781	6.0070
	3	3.2448	5.3405	4.1594	5.6913
	4	1.9118	3.0852	3.5924	5.0549
	5	4.7474	6.2010	5.5186	8.1848
kebab-case	1	1.5608	3.9041	5.7515	9.4349
	2	1.4810	4.1854	5.2516	3.9117
	3	1.7487	1.9074	6.4289	2.8454
	4	2.6956	2.1648	4.6879	4.7700
	5	2.8275	4.8684	5.9331	4.7426

Here we plot the results for every participant in order to better understand what are the numbers in the table:
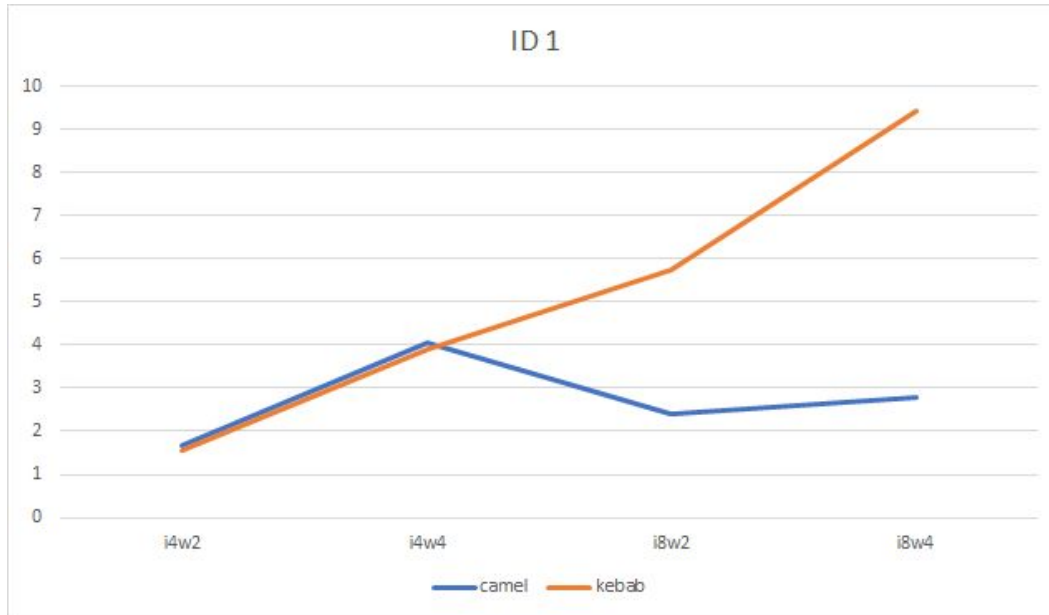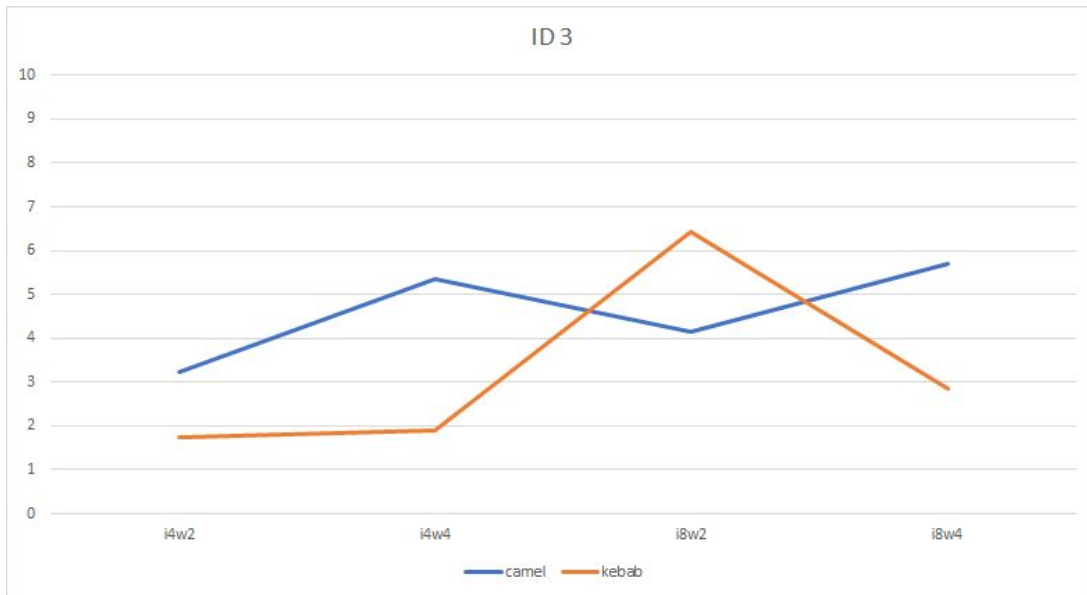

Here we can see the measurements of the first participant. We can see that with 4 identifiers the kebab-case style seems to be better, when using 8 identifiers the camelCase style is not only better than the kebab-case style, but also the participant performed better in comparison with the 4 identifiers.
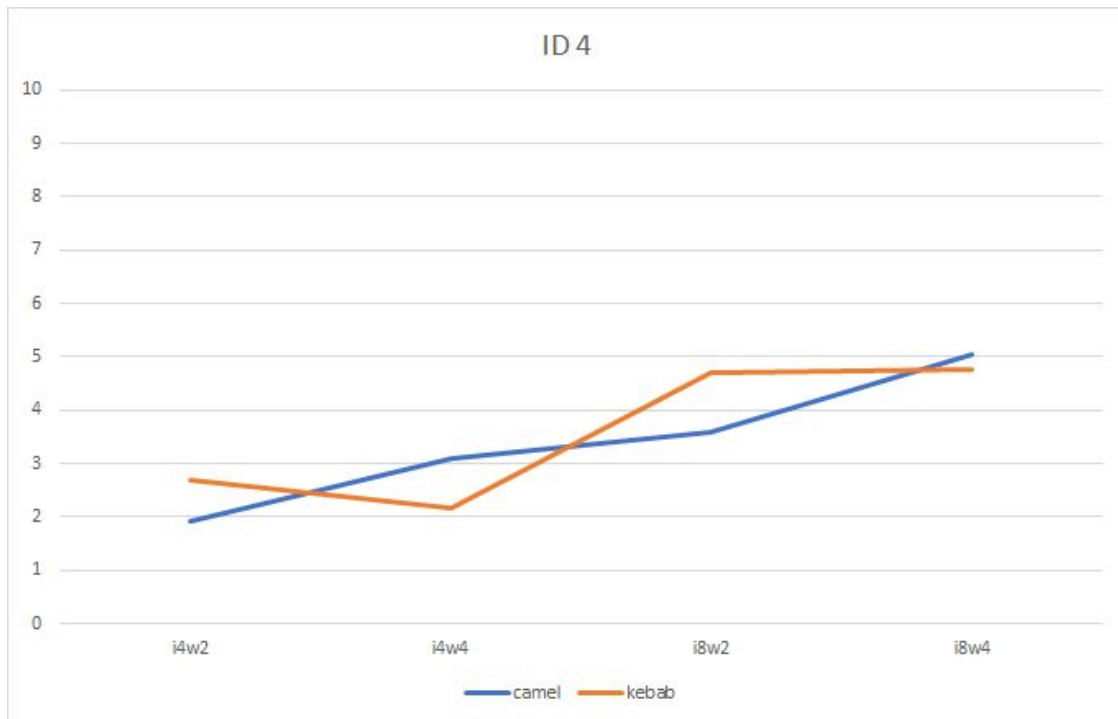
Here we can see the measurements of the second participant. We can see how the kebab-case gave some advantage to the participant as she performed better in every case.
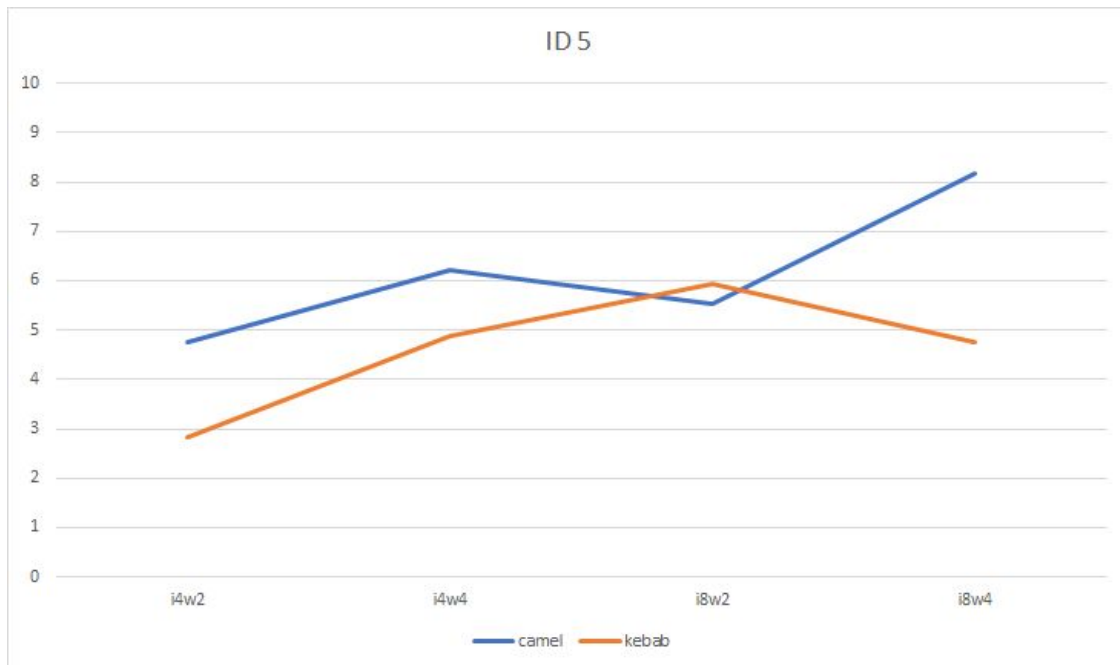

Here we can see the measurements of the third participant. We can see how the kebab-case gave some advantage to the participant as he performed better in almost every case, only when confronted with 8 identifiers and with an identifier with 2 word he performed worse.

Here we can see the measurements of the fourth participant. We can see that when using words with 2 identifiers, the camelCase style gave an advantage to the participant. When using 4 words the kebab-case had a better impact on the performance.

Here we can see the measurements of the fifth participant. We can see how the kebab-case gave some advantage to the participant as he performed better in almost every case, only when confronted with 8 identifiers and with an identifier with 2 words he performed worse.

Experiment Report Template — E&E'20 — Stefano Gonçalves Simao

Here I report the accuracy (number of errors):

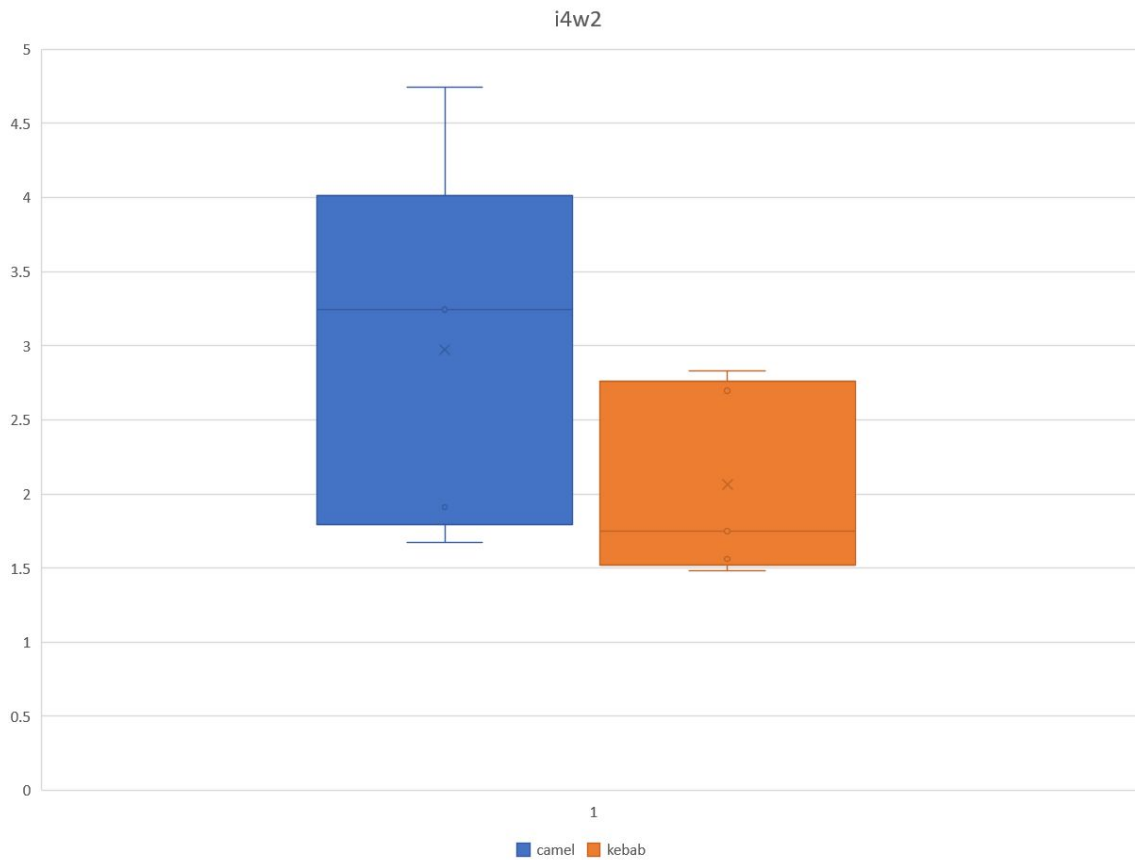| Style	ID	4 identifiers 2 words	4 identifiers 4 words	8 identifiers 2 words	8 identifiers 4 words	Total
camel	1	1	0	0	1	2
	2	0	0	0	0	0
	3	1	0	0	2	3
	4	0	0	1	1	2
	5	1	0	1	2	4
kebab	1	0	0	0	0	0
	2	0	1	0	0	1
	3	0	0	0	0	0
	4	1	0	0	1	2
	5	0	0	1	1	2

From the accuracy we can see the camelCase produced 11 errors and the kebab-case only 5. We can also see that the participant 5 made more mistakes than the others.
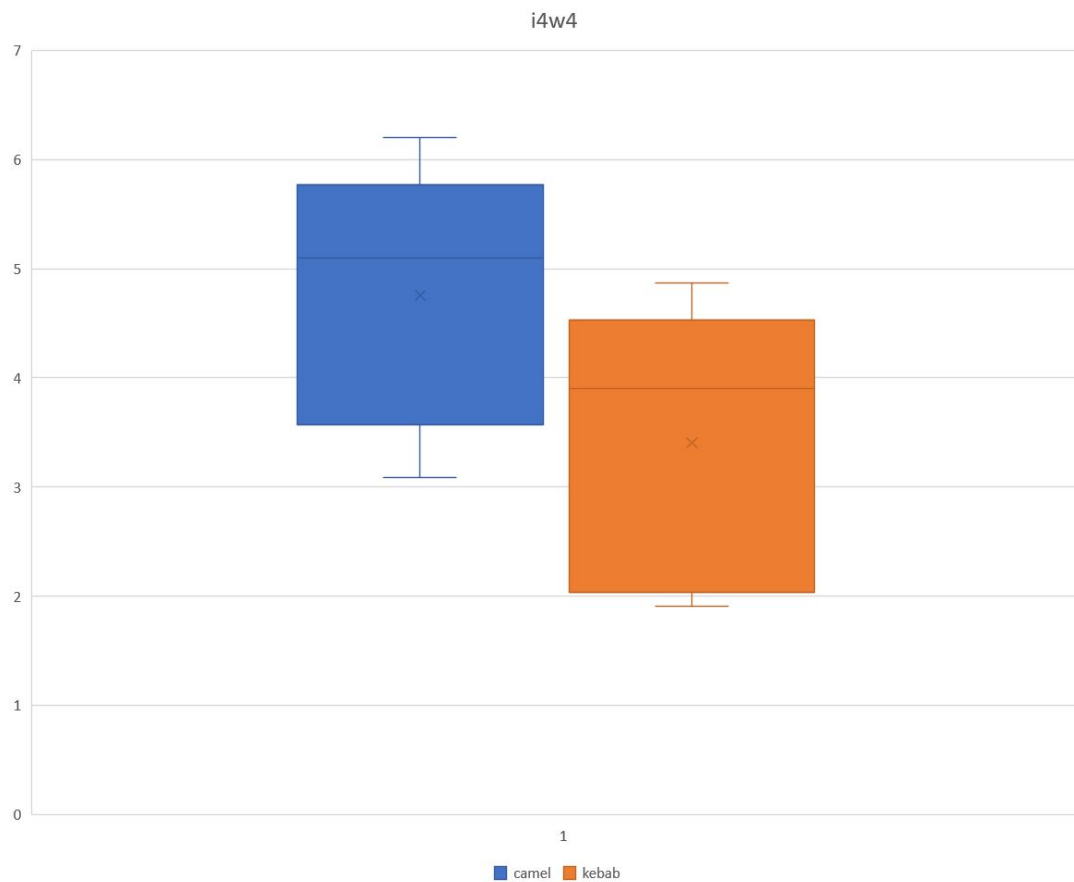
3.2 Descriptive Statistics

In this table I report the statistics of the experiment. In the following box-plots I visualize the contents of this table, where we can really see what are the results of the experiment.

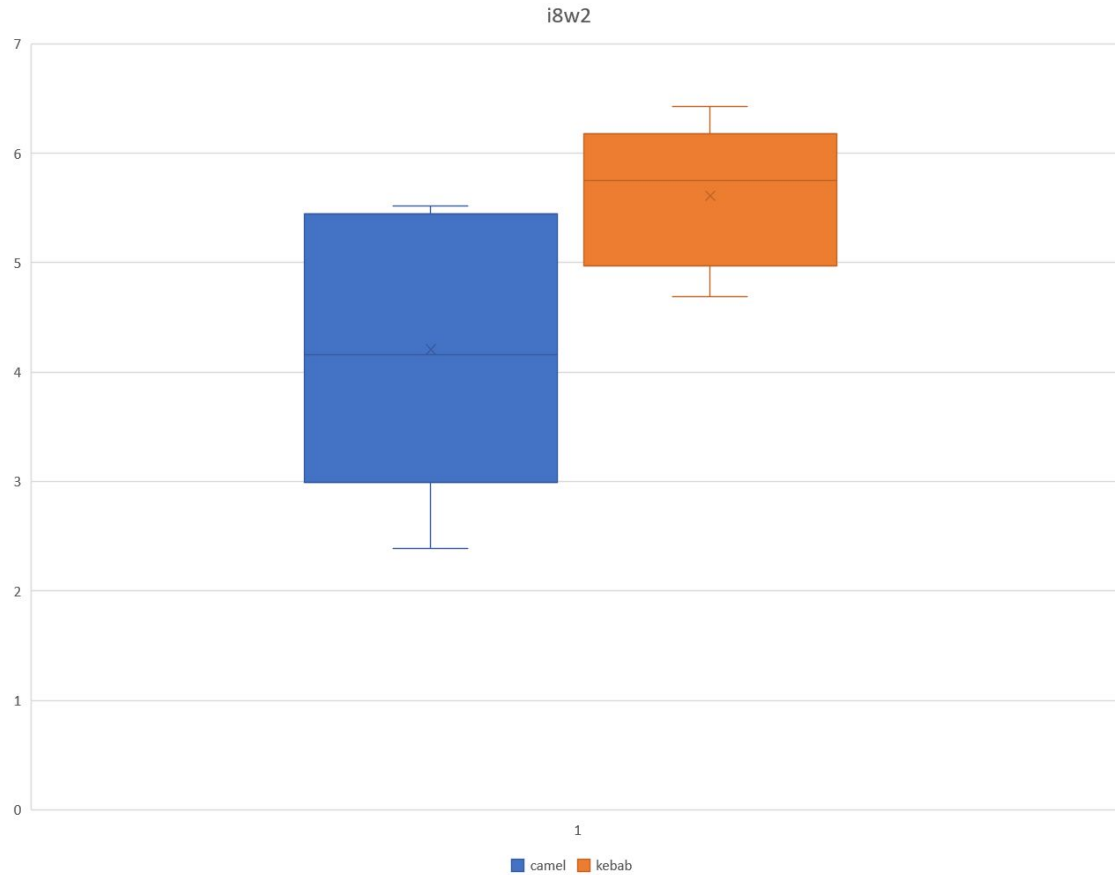| Style	Statistics	4 identifiers 2 words	4 identifiers 4 words	8 identifiers 2 words	8 identifiers 4 words
camelCase	Minimum	1.6756	3.0852	2.3853	2.7815
	1st Quartile	1.9118	4.0604	3.5924	5.0549
	Median	3.2448	5.0966	4.1594	5.6913
	3rd Quartile	3.2755	5.3405	5.3781	6.0070
	Maximum	4.7474	6.2010	5.5186	8.1848
kebab-case	Minimum	1.4810	1.9074	4.6879	2.8454
	1st Quartile	1.5608	2.1648	5.2516	3.9117
	Median	1.7487	3.9041	5.7515	4.7427
	3rd Quartile	2.6956	4.1855	5.9332	4.7700
	Maximum	2.8275	4.8684	6.4289	9.4349

i4w2

In this box plot for the case with 4 identifiers and 2 words we can clearly see that the kebab-case style produced better results in this experiment. The median, minimum and maximum are lower than the camelCase style.
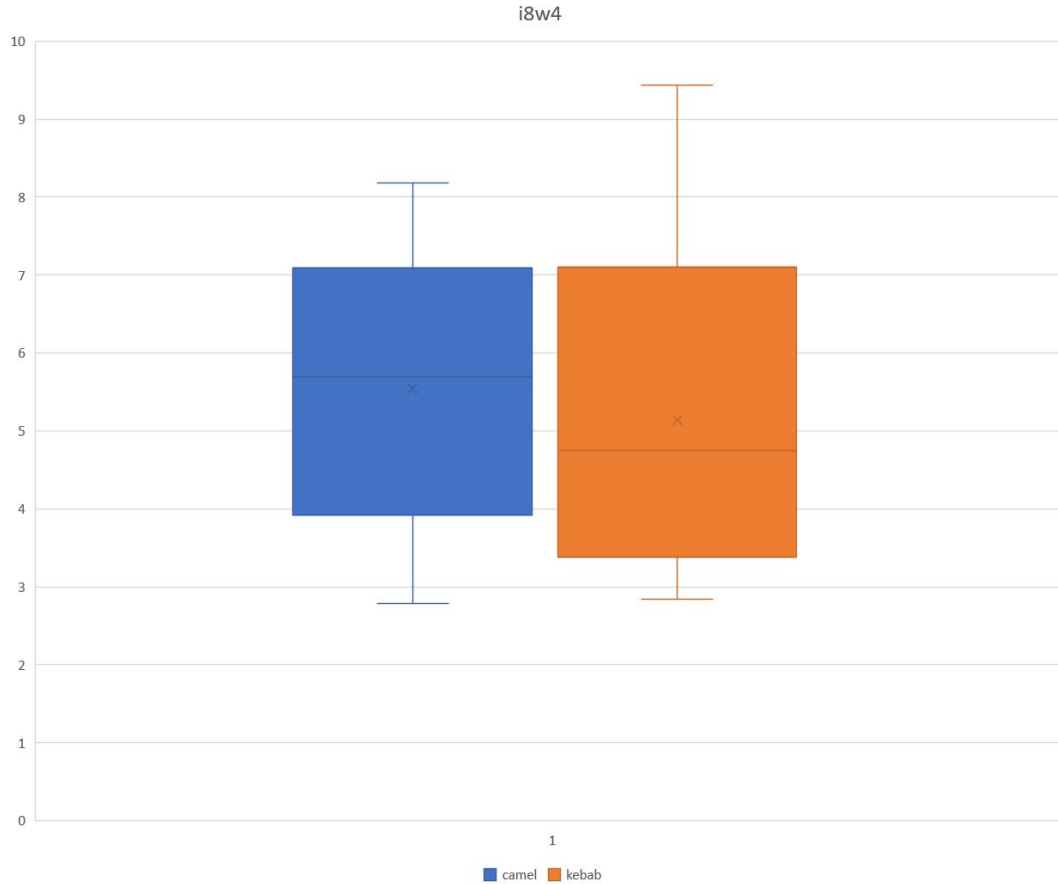
i4w4

camel kebab

In this box plot for the case with 4 identifiers and 4 words we can clearly see that the kebab-case style produced better results in this experiment. The median, minimum and maximum are lower than the camelCase style.

i8w2

camel   kebab

In this box plot for the case with 8 identifiers and 2 words we can clearly see that the camelCase style produced better results in this experiment. The median, minimum and maximum are lower than the kebab-case style.

i8w4

In this box plot for the case with 8 identifiers and 4 words we can see that the kebab-case style produced better results in this experiment. The kebab-case had a maximum of 9.4349, that is bigger than the maximum of the camelCase, but the median is lower. It also has a bigger interquartile range because it has a bigger standard deviation, as we can see in this table:

Style	Statistics	4 identifiers 2 words	4 identifiers 4 words	8 identifiers 2 words	8 identifiers 4 words
camelCase	Mean	2.971068	4.756764	4.206799	5.543966
	Standard Deviation	1.237229	1.206336	1.30291	1.941694
kebab-case	Mean	2.062769	3.406067	5.610652	5.140964
	Standard Deviation	0.646949	1.301995	0.665893	2.525603

Also from this table we can see that the mean shows a better performance from the kebab-case style in 3 of the 4 experiments (4 identifiers and 2 words, 4 identifiers and 4 words, 8 identifiers and 4 words). The standard deviations for the camelCase are consistent but for the kebab-case

I had bigger variations but also a lot more concentrated results (4 identifiers and 2 words, 8 identifiers and 2 words.

4. Discussion

4.1 Compare Hypothesis to Results

The results show that the kebab-case style offers better readability as, except for the case 8 identifiers and 2 words, the participants have been able to find the correct identifier faster than with the camelCase. If we look at every participant we can see that the kebab-case style is not faster in every test. In some cases it is noticeably slower, for example for the first participant when dealing with 8 identifiers. These differences can be attributed to the position of the correct identifier not being where the eyes of the participant are when the words show up on the screen, and also being at the end of the pattern chosen to read the identifiers. In fact when the camelCase obtained a better result is when there are 8 identifiers on the screen, pointing to the fact that it may be for this reason and not for the fact that camelCase is better.
We can also see that the last participant has consistently been the worst or almost worse performer. If we look at the others we can conclude that the English level seems to have an impact in the speed of reading but not in determining which style is better, as also here the kebab-case is better on average.
If we look at the errors made by the participants, these also show how the camelCase seems to produce more errors as the brain may interpret the identifier as one big word and reads only part of it filling the gaps. It is very well known that if we omit some letter from a text, our brain is able to read it with no problem. Furthermore, if we introduce some errors, our brain may very well not be able to see them when reading. This also can impact the kebab-case but having the word separated may push our brain to read more parts of the words, i.e. more letters, and hence reducing the errors.
The statistical analysis of this experiment proves my hypothesis.

4.2 Limitations and Threats to Validity

Internal validity could be threatened by the fact that the participants may get better after the first test (i.e. when tested with 4 identifiers and 2 words) because after that they try harder to give the correct answer in faster times. Another threat is related to the small size of the number of participants, with only 5 participants we can't have a meaningful conclusion but only an indication on which style may be better. For example, as discussed above, a participant may be lucky and, when the test starts, may be already looking at the correct word. Or maybe the participant is more focused at the time of the experiment, or maybe it didn't sleep enough that night and this can have an impact on the cognitive performance that particular day.

External validity can be jeopardized by the fact that these results could be attributable to how the experiment was made and is not generalizable to real world use cases. For example, the camelCase style could be better suited for smaller font sizes or when used in a programming context.

4.3 Conclusions

The results of this experiment clearly indicates that the kebab-case style gives faster readability and the participants made fewer mistakes. If we look at the descriptive statistics it is clear that the camelCase style entails more effort when reading and it is more prone to mistakes.

Appendix

A. Materials

The only material used for this experiment is the program used to measure the times. I included all the code used for reproducibility purposes.

B. Reproduction Package (or: Raw Data)

The raw data excel sheet is included in the file "EE2_raw_results.xlsx". Here there are the measurements and computation with full precision.