camelCase vs kebab-case

Experiment 2, Experimentation & Evaluation 2024

# Abstract

This study investigates whether identifier style, camelCase or kebab-case, affects readability in terms of speed and accuracy. Using a web application developed in VUE, participants performed a recognition task involving sentences written in both styles. Each participant completed 20 tasks, 10 for each style, selecting the correct identifier from a set of options. Response times and accuracy data were collected and stored in a CSV file for analysis, which were then run through JASP in order to gather statistical data used for creating boxplots and independent t-tests. These were then employed to compare the readability of the two styles. The findings aim to provide insights into optimal coding practices for improving code readability and developer efficiency.

# 1. Introduction

In the realm of software development, readability plays a crucial role in maintaining and understanding source code. Composed identifiers, meaning names consisting of multiple words, are omnipresent in programming, as they help convey the purpose of variables, functions, and classes. However, the way these identifiers are formatted may impact a developer's ability to read and understand them efficiently. While studies in natural language have shown that explicit separators (e.g., spaces or symbols) improve reading speed and comprehension, it remains unclear whether these findings extend to programming languages, where composed identifiers often follow conventions like camelCase or kebab-case.

In this study, we aim to explore whether identifier style affects the speed and accuracy of recognition. Specifically, we compare camelCase, which uses capital letters to separate words, to kebab-case, which employs hyphens as separators. Understanding the implications of these styles can guide best practices for naming conventions and improve code readability, ultimately enhancing developer productivity.

To investigate this, we designed a controlled experiment where participants identified correct identifiers written in camelCase or kebab-case from a set of distractors. By analyzing response times and accuracy, we seek to determine whether one style is inherently more readable than the other.

| **Hypotheses:** |
| --- |
| Hypotheses 0: No matter the identifier style, both camelCase and kebab-case  Main hypotheses: camelCase is faster to read rather than kebab-case  Write down your (falsifiable!) hypotheses here. Each hypothesis must include **independent** and your **dependent** variables. You must write down your hypotheses **before** you do your experiment! |

# 2. Method

In the following subsections, describe everything that a reader would need to replicate your experiment in all important details.

## 2.1 Variables

Explicitly identify the independent variable(s) (i.e., what you as the experimenter manipulate):

| **Independent variable** | **Levels** | |
| --- | --- | --- |
| Identifier styles | camelCase or kebab-case | |

Explicitly identify the dependent variable(s) (i.e., what you measure):

| **Dependent variable** | **Measurement Scale** | |
| --- | --- | --- |
| Reading speed  Reading accuracy | Time measured in seconds correct/incorrect response | |

Explicitly identify any important control variable(s) (i.e., what you keep constant): Note that you do *not* need to spell out items that you do not expect to make a *significant* difference! E.g., do not list room temperature unless you believe that minor differences have an impact! Or the other way around: only list variables here that you think are important to keep at a certain level.

| **Control variable** | **Fixed Value** | |
| --- | --- | --- |
| Sentences | Difficulty, number of sentences used | |

Explicitly identify the blocking variable(s) (i.e., potential sources of variability you measure and will use to partition the experimental units into blocks, but that are not part of the hypothesis):

| **Blocking variable** | **Levels** | |
| --- | --- | --- |
| TODO |  | |

## 2.2 Design

Check off the characteristics of your experimental 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): [for human subject studies]

| ⃞ **Between Group Design** (independent measures) | ⃞ **Within Subject Design** (repeated measures) | ⃞ Other |
| --- | --- | --- |

Explain, (1) *in text using terminology from the book and lectures* **and** (2) with a figure (similar to those used in Chapter 3 of the Field & Hole book), what kind of experiment you did.

## 2.3 Participants

Describe who will take / took part in your experiment. Provide descriptive/summative statistics of their gender, age, professional backgrounds, and any other characteristics that may be relevant to your experiment. Also explain how you will recruit / recruited them (volunteers recruited through email, classmates who were asked to do this, etc) and how you will allocate / allocated them into the different study conditions, i.e., control group vs experimental group(s).

Participants were taken from a group of people that included people with mostly some programming experience, but only a small number that had none. The age range was in the early to mid 20’s, with only one being over that. There was no requirement for specific expertise in camelCase or kebab-case, which allowed for a mix of familiarity levels. To examine possible effects on performance, demographic information such as age, years of programming experience, and preferred coding style was gathered. As for control group

## 2.4 Apparatus and Materials

Describe in sufficient detail any relevant “props” that you used in your experiment. This could be the computer you used (exact model and specification), the software used (URL, version numbers), the way you measured, e.g., time (A stopwatch? A background process on the computer that got automatically triggered?). Omit needless detail (e.g., think whether details like the size of the table the laptop was placed on, or the hard disk size, might have affected your results or not).

The experiment was done by using a web application developed through the VUE framework and Javascript. Its main feature was the recording of the 20 answers from the participants, divided into 10 camelCase and 10 kebab-case experiment tasks, which were then saved into a .csv file, containing all the information needed to perform data analysis through JASP, mainly the time taken for the participant to click on the answer and whether or not it was the correct one. The application also contained a page instruction to explain how the test worked, as well as demographics form to gather data for programming/work experience, age and identifier/programming language preferences.

## 2.5 Procedure

Describe how you used your props and the participants to perform your actual experiment, i.e., how you actually carried out a single experimental run. What was done to the participants? What did they have to do? How long did each session take (unless this is an actual dependent variable)? If you did not have participants, explain, e.g., what software was started by whom in what order.

The procedure was divided into 4 parts:

**Introduction and Consent:** Participants were welcomed, provided with instructions, and gave consent to participate.

**Demographics Collection:** Participants completed a short form capturing key demographic and professional data.

**Task Execution:** Participants were shown a target sentence followed by a set of identifiers. Their task was to identify the correct match as quickly and accurately as possible.Each participant completed 20 trials, with camelCase and kebab-case tasks randomized to minimize learning effects.

**Data Collection:** Response times and accuracy for each task were automatically recorded and saved for analysis.

# 3. Results

## 3.1 Visual Overview

Provide an insightful overview of the data you collected. This requires some engineering from your part, to find a good degree of summarization: On one end of the spectrum, you don't summarize, and report hundreds of raw measurement values in a block of text. On the other end of the spectrum, you report a single number (like a mean value). Both approaches are bad.

Instead, use appropriate visual summaries (such as **scatter plots**, **histograms**, **box plots**, or **empirical cumulative distribution functions**) to show the distribution of your data. If you have a very small number of measurement values, then report all of them in a **well organized table** (where rows and/or columns correspond to different levels of different factors).

## 3.2 Descriptive Statistics

For each group or condition, summarize the set of measured values with a "five-number summary": **minimum**, **first quartile**, **median**, **third quartile**, and **maximum** (note: these are the statistics underlying a box plot).

Moreover, report the **mean** and **standard deviation** (note: for data that is not normally distributed, e.g., for multi-modal data, these two statistics may be less meaningful).

Make sure you explain – in your words – what these statistics mean “in plain English”, but don’t yet interpret them (this is for the Discussion section).

As a preface, we gathered 200 results from 10 subjects, but only 95% of them were used, since we do not consider incorrect answers in our statistical calculations. Of the incorrect answers, 7% were from the kebab-case, while the remaining 3% camelCase. This can be taken as a pre-analysis sign that camelCase is slightly easier to read than kebab-case.  
Looking at the descriptive data, we’ve come to find out that the results are evenly split in terms of identifier performances and in only 2 cases the preferred identifier didn’t match with the mean obtained.  
  
Putting all the results together, we obtain that the median of camelCase is 3.850s, while for kebab-case 3.520s, which is slightly lower. The mean times are also close, with camelCase (4.315 s) slightly higher than kebab-case (4.282 s). However, the difference in means is very minimal, suggesting no significant advantage for either style on average.

Standard deviation is smaller for camelCase, suggesting that participants' performance was more consistent with camelCase compared to kebab-case, which shows greater variability.

The minimum and maximum times reveal a wider range for kebab-case (1.300 s to 18.180 s) compared to camelCase (1.520 s to 12.150 s). This variability could be influenced by outliers or differences in participants' familiarity with kebab-case.

In conclusion, the data obtained indicates that both cases have similar average and median processing times, with camelCase showing slightly more consistent performance across participants. The null hypothesis is at this moment confirmed, contradicting our main assumption.

## 3.3 Inferential Statistics

If applicable, you then follow these up with inferential statistics – i.e., the **results of statistical tests** that you did in order to decide whether there were any “real” (i.e., not by chance) differences between the conditions/groups. You should also explain what statistical test you used, and, if not immediately obvious, why.

Make sure you explain – in your words – what these statistics mean “in plain English”, but don’t yet interpret them (this is for the Discussion section).

By inferential statistics we mean that with the data obtained in the previous section, we make some calculations to help with the conclusion drawing process of our data sample. In this case, we used the Independent Samples T-test, useful in cases like this where the amount of data for both groups isn’t equal.

With a t-statistic of 0.099, we see that there’s a small difference between the means of each identifier’s result and with a p value of 0.921 (which is greater than 0.05), we confirm what we have observed earlier through the descriptive statistics, which is that there’s no statistically significant difference in the time taken between camelCase and kebab-case. A Cohen’s d of 0.014 suggests also that there’s a very small or negligible effect size.

# 4. Discussion

## 4.1 Compare Hypothesis to Results

Provide a brief restatement of the main results from the previous section, and if (or if not) these support your research hypothesis.

If there is a discrepancy between your hypothesis and the results of your experiment, speculate about why you were unable to find evidence to support your hypothesis.

To summarize what we have obtained through our statistics, we found out that our null hypothesis was not disproven, and therefore we can say that there’s no discernible advantage in time between reading camelCase and kebab-case. The only thing that is sure is that camelCase seems more stable compared to kebab-case, which is proven by the amount of subjects that chose it as their preferred identifier compared to kebab-case.

## 4.2 Limitations and Threats to Validity

Acknowledge any limitations and threats to validity of your study, and how seriously these affect your results. How could these be remedied in future work?

## 4.3 Conclusions

End with the main conclusions that can be drawn from your study.

Appendix

# A. Materials

Any documents you used for your informed consent (information sheets, consent) or as part of your apparatus (e.g., manual, hand-out), please include them here.

# B. Reproduction Package (or: Raw Data)

Before, during, and after the experiment you collected all kinds of data. Don't ever throw such data away! Any plots, tables, summaries, and statistics provided in this report should be recreatable from the raw data you have.

If you only collected a small amount of data, put it in this Appendix right here.

If you collected data in forms that are better kept in separate files, then zip up those files, and submit them as a "reproduction package" supporting this report.