Replication Study: Eye Tracking Study on camelCase and under_score Identifier Styles

Johannes Almroth

27th May 2019

Abstract

The purpose of this paper is to replicate the results of two earlier studies, Binkley *et al.* (2009) as well as Sharif *et al.* (2010), where the aim was to determine if the identifier naming conventions affected code comprehension. The two styles that are examined are camelCasing as well as the under_score style. The study is conducted by using an eye-tracking device to gather data through a timed responsed test.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

- Identifier names are important

- We have two identifier styles

- Some history

- Two other studies

- Some background on them

- Reference all terms introduced

1.1 Purpose and Goals

- Write about the grand scheme of things

- Set the correct expectations

- What can I expect to learn if I keep on reading?

- What are the success criteria for this work?

- How will the work be evaluated?

1.2 Delimitations

- Scale down expectations and clarify

I will limit the project by not extending the original study with more than a few additional variables that are relevant to the subject. I will furthermore limit the size of participants to the minimum amount necessary to achieve statistical significancy, if time doesn't permit otherwise. All data gathered will solely be through the same means used in the original 2010 study, the tobii eye tracker as well as interviews.

Chapter 2

Background

- Have a specific background or spread it out

- May be necessary to introduce certain things

- From our running example, we might discuss prior work on simplifying parallel programming, or SOTA with e.g., threads

In 2010 a study published at ICPC examined how identifier naming conventions affect code comprehension with the help of eye tracking equipment. The two styles that were analysed were camelCase as well as under_score identifiers. The goal of this project is to try and replicate the results of the study with the same method, as well as taking more variables into account, if time and the sample size of the study group permits.

Identifier names are keystones in software programming for presenting and working with data in any kind of function. If any way of writing these identifiers improves the speed of which we comprehend the code, it could potentially entail better overall program understanding.

Historically, underscores were used as the primary identifying style because of early programming languages being case insensitive. Later on, camel-case identifiers became the norm, maybe partially because of the number of keystrokes and ease of writing it had over using underscores.

The results from the 2010 study concluded that there were no significant differences in regards to accuracy between the two styles. However, the camel-cased identifiers took longer to identify. The 2010 study in turn sought to replicate a study conducted in 2009, which similarly concluded that camel-cased identifiers were slower to identify, but that underscored identifiers brought about less accuracy. It's worth mentioning that the two studies differed their methods of gathering data, as the 2009 study used a game like interface to gather timed responses, while the 2010 study used eye tracking equipment to gather more quantitative data.

Chapter 3

Methodology

- How to measure simplicity and scalability

- Not just what you did but why and what the consequences are

3.1   Research questions

Explore design space, motivate method, consequences,

I will attempt to replicate the 2010 study using an eye tracker device from tobii. I will set up a test process as similar to the original as I can, as to gather a minimum of the same amount of data. My aim is also to gather more data about the subjects that could act as confounding variables, such as what programming languages that they've used historically and currently, age gap, and other variables TBD (to be determined).

In the original test process subjects were instructed to observe an identifier on a screen, which was then replaced with another four identifiers. One of them matched the previous identifier, whereas the rest of them were incorrect, meant ot distract the subject. The data that was then measured is the accuracy of the subject to identify the correct word, as well as the time it took to reach an answer. From the tobii eye tracker additional data such as time spent on each word, amount of fixations vs saccades (eye flicker), can be extracted. Furthermore, data such as subject age, experience and background will be extracted in an interview like setting.

3.2   Eye-tracking Equipment

3.3   Material and Stimuli

3.4   Visual Effort and Areas of Interest

3.5   Study Variables

3.6   Hypotheses

3.7   Participants

3.8   Instrumentation

Chapter 4

Results

- Don't make the reader do all the work

- Have hypothesis, test them, state result clearly

- Two lists are not a comparison

- Be the first to criticise your own work

4.1 Correctness and Find Time

4.2 Visual Effort

4.3 Simularities and Differences

Chapter 5

Discussion

- Don't make the reader do all the work

- Have hypothesis, test them, state result clearly

- Two lists are not a comparison

- Be the first to criticise your own work

Chapter 6

Threats to Validity

Chapter 7

Conclusions & Future Work

- Critique

- Discussion

- Don't make the reader do all the work

7.1   Future Work

Bibliography