A Replication on the Effect of Identifier Style on Comprehension

Bonita Sharif
University of Nebraska – Lincoln
Lincoln, Nebraska, USA 68588

Jonathan I. Maletic
Kent State University
Kent, Ohio, USA 44242

WHO

Original Authors: D. Binkley, M. Davis, D. Lawrie and C. Morrell, (2009), "To camelcase or under_score", *17th International Conference on Program Comprehension (ICPC 2009)*, Vancouver, BC, 2009, pp. 158-167. doi:10.1109/ICPC.2009.5090039

Authors Performing Replication: B. Sharif and J. I. Maletic, (2010), "An Eye Tracking Study on camelCase and under_score Identifier Styles",*18th International Conference on Program Comprehension (ICPC 2010)*, Braga, Portugal, 2010, pp. 196-205. doi:10.1109/ICPC.2010.41

Joint Journal Submission: Binkley, D., Davis, M., Lawrie, D., Maletic, J., Morrell, C., and Sharif. B., (2013), "The Impact of Identifier Style on Effort and Comprehension" *invited submission* to the Springer *Journal of Empirical Software Engineering, vol 18. Issue 2*, pp. 219-276.

WHAT

The study done by Binkley et al. was replicated by Sharif et al. in 2010 using a different method of data collection namely an eye tracker with a different set of human participants. Jointly, the two teams then wrote up results of two additional experiments on the same topic and published this work in the EMSE journal two years later in 2013.

WHY

Identifiers comprise of nearly 70% of source code. If a certain identifier naming style significantly speeds up program comprehension this could significantly impact overall program understanding. Results could also become evidence for code style and readability guides.

HOW

The initial paper by Binkley et al. studied the effect identifier style has on correctness and efficiency. The data was collected via an online questionnaire where participants were asked to click on identifiers that were generated in either camel case or underscore from a 2-word or 3-word phrase.

WHERE

The replication done a year later used the same 2-word and 3-word phrases but used another method of data collection. Instead of using an online questionnaire, an eye tracker was used to collect developer gaze as they answered the question of choosing the correct identifier formed from phrases. The replication used same words and phrases from the Binkley et al. paper. However, additional metrics such as visual effort were used in the replication.

DISCUSSION

The replication was done in a completely different setting than the original study even though the tasks were the same. The experiment protocol needed to be adjusted to account for the eye tracking equipment. One of the lessons we learned was multi-institutional replications such as the one we undertook made us understand that the population being tested plays a big role in your results. For example, one result we found is that people who were trained in the underscore style were better at that style than those who were trained in the camel case style. The end result from the family of experiments we undertook showed camel case performing better. The results of this study were used as evidence to support the choice of camel casing in languages such as Quorum by Stefik et al. This shows impact of such replications.