Women and Men – Different but Equal: On the Impact of Identifier Style on Source Code Reading

Zohreh Sharafi, Zéphyrin Soh, Yann-Gaël Guéhéneuc, and Giuliano Antoniol

Ptidej Team and Soccer Lab, École Polytechnique de Montréal, Montréal, Canada.
Program evolution is often solely based on source code.

Program comprehension heavily depends on identifiers.
Introduction

- What is the impact of identifiers on program comprehension?
  - Identifier Style (Sharif 2010\(^1\), Binkley 2009\(^2\))
    - Camel Case
    - Underscore
  - Identifier quality (Lawrie 2007\(^3\))
    - Length
    - Structure

---

1. D. Binkley, M. Davis, D. Lawrie, and C. Morrell, “To camelcase or under score,” ICPC, 2009
# Identifier Style

<table>
<thead>
<tr>
<th>Investigate the Impact of Identifier Style</th>
<th>“Camel Case vs. Underscore”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
</tr>
<tr>
<td>Binkley, 2009¹</td>
<td>Sharif, 2010²</td>
</tr>
<tr>
<td>Higher accuracy for Camel Case</td>
<td>No difference</td>
</tr>
<tr>
<td><strong>Time</strong></td>
<td></td>
</tr>
<tr>
<td>More time spent on Camel Case</td>
<td>More time spent on Camel Case</td>
</tr>
<tr>
<td><strong>Effort</strong></td>
<td></td>
</tr>
<tr>
<td>Not applicable</td>
<td>Lower visual effort on Underscore</td>
</tr>
</tbody>
</table>

1. D. Binkley, M. Davis, D. Lawrie, and C. Morrell, “To camelcase or underscore,” CPC, 2009
Gender Preference

- Does gender affects identifier style preference?

- Does gender and identifier style affects program understanding?
## Related Work

<table>
<thead>
<tr>
<th>Self-efficacy (1994)¹</th>
<th>Lower self-efficacy</th>
<th>Higher self-efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selectivity Hypothesis (2001)²</td>
<td>Use comprehensive way</td>
<td>Use simple heuristic</td>
</tr>
<tr>
<td>Software Adaptation &amp; Use (2003)³</td>
<td>No experience = No self-confidence</td>
<td>Apply general knowledge for specific tasks</td>
</tr>
<tr>
<td>Testing Strategies (2009)⁴</td>
<td>Use testing only for finding bugs</td>
<td>Use testing for: 1. finding bugs 2. fixing bugs 3. evaluating their fix</td>
</tr>
</tbody>
</table>

---

Our Empirical Study

Goal: Design and perform an experiment to investigate the impact of gender on identifier styles and identifier comprehension strategies

Research question: Does the developers’ gender impact their effort, their required time, and as well as their ability to recall identifiers (accuracy) in source code reading?

- Memorability impacts the comprehension of identifiers\(^1,2,3\)
  - Lacking memorability may impair:
    - Identifiers recall ➔ Code readability ➔ Program comprehension

---

2. B. Sharif and J. Maletic, “An eye tracking study on camelcase and underscore identifier styles, ICPC, 2010
Experiment Design

- **Independent variables**
  1. Gender: male (M) or female (F)
  2. Identifier style: Camel Case (CC) or Underscore (US)

- **Dependent variables**
  1. Accuracy
  2. Required time (Speed)
  3. Visual effort

- **Mitigating variables**
  1. Study level
  2. Style preference of subjects
Experiment Design

Subjects’ Demography
(24 Subjects)

<table>
<thead>
<tr>
<th>Academic background</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ph.D.</td>
<td>M.Sc.</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
</tr>
</tbody>
</table>

Source code stimulus

```java
import java.util.Scanner;

public class PrimeNumberCalc {
    private final int UpperLimit = 100;
    public void calcPrimeNums() {
        int nCounter = 0;
        int i = 0;
        boolean isPrimeNumber = false;

        while (++i <= UpperLimit) {
            int innerLoop = (int) Math.ceil(Math.sqrt(i));
            isPrimeNumber = false;

            while (innerLoop > 0) {
                if ((i != innerLoop) && (i % innerLoop == 0)) {
                    isPrimeNumber = false;
                    break;
                } else if (!isPrimeNumber) {
                    isPrimeNumber = true;
                } else {
                    --innerLoop;
                }
            }
            if (isPrimeNumber) {
                System.out.println(i);
                ++nCounter;
            }
        }
        System.out.println("Nr of prime numbers found: " + nCounter);
    }
}
```

What is the output of this program?

Question stimulus

Q1: what is the name of the class?

1. PrimeNumClac
2. NumCalculator
3. myClass
4. PrClass
Experiment Design
Eye-tracking system

- **FaceLAB**: 
  - Video-based
  - Two cameras
  - An infrared sensor

- Non-intrusive
  - No goggles,
  - No wires
  - No sensing device

Results and Analyses: Accuracy and Time

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camel Case</td>
<td>83%</td>
<td>80%</td>
</tr>
<tr>
<td>Underscore</td>
<td>68%</td>
<td>85%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camel Case</td>
<td>5.25</td>
<td>6.6</td>
</tr>
<tr>
<td>Underscore</td>
<td>6.7</td>
<td>7.8</td>
</tr>
</tbody>
</table>

Table 1 - Accuracy

Table 2 - Time

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>74%</td>
</tr>
<tr>
<td>Women</td>
<td>82%</td>
</tr>
</tbody>
</table>

Table 3 - Overall Time and Accuracy

No significant differences

Accuracy

No significant differences

Time
Results and Analyses: Visual Effort

- Visual effort:
  - Calculated from eye-tracking data
  - Calculated based on the amount of visual attention:
    - Less attention $\rightarrow$ Less effort

- Visual attention can trigger the mental processes

- Two types of eye gaze data:
  - Fixation
  - Saccade

- We use fixation to calculate effort as previous works
Results and Analyses: Visual Effort Metric - Source code Stimulus

- Convex hull: the smallest convex sets of fixations that contains all of a subject’s fixations\(^1\)

- Smaller convex hull
- Close fixations
- Less effort

---

Results and Analyses:
Visual Effort Metrics - Question Stimulus

Q1) What is the name of the method used for executing the query?

1. executeQuery
2. myMethod
3. runFunction
4. dBMethod

Metrics to measure visual effort (Sharif 2010\textsuperscript{1})

<table>
<thead>
<tr>
<th>Numbers of eye fixations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$FC(Q) = \sum_{a \in \text{tasks, all answers}} f(a)$</td>
</tr>
</tbody>
</table>

| $FR(\text{correct}) = \frac{\sum_{a \in \text{correct answer}} f(a)}{\sum_{a \in \text{correct answer} \cup \text{distracters}} f(a)}$ |
| $FR(\text{distracters}) = \frac{\sum_{a \in \text{distracters}} f(a)}{\sum_{a \in \text{correct answer} \cup \text{distracters}} f(a)}$ |

---

I. B. Sharif and J. Maletic, “An eye tracking study on camelcase and under score identifier styles,” IEEE 18th International Conference on Program Comprehension, 2010
### Result and Analysis: Visual Effort

<table>
<thead>
<tr>
<th>Effort</th>
<th>FC(Q)</th>
<th>FR(Correct)</th>
<th>FR(distracter)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No significant differences</td>
<td>No significant differences</td>
<td>Female subjects spends more visual effort</td>
</tr>
</tbody>
</table>

- Female subjects put more visual effort on distracters
- Female subjects analyze all distracters before making decision
Results and Analyses: Visual Effort

Heatmap

(a) Q4) what is the name of the class?
1) 2D_frame
2) Painter
3) Java_2D_frame
4) My_class

(b) Q4) what is the name of the class?
1) 2D_Frame
2) Painter
3) Java_2D_Frame
4) My_Class
## Conclusion

### Investigate the Impact of Identifier Style

“Camel Case vs. Underscore”

<table>
<thead>
<tr>
<th></th>
<th>Binkley, 2009(^1) (135 subjects)</th>
<th>Sharif, 2010(^2) (15 subjects)</th>
<th>Our Experiment (24 subjects)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>Higher accuracy for Camel Case</td>
<td>No difference</td>
<td>No difference</td>
</tr>
<tr>
<td><strong>Time</strong></td>
<td>More time spent on Camel Case</td>
<td>More time spent on Camel Case</td>
<td>No difference</td>
</tr>
<tr>
<td><strong>Effort</strong></td>
<td>Not applicable</td>
<td>Lower visual effort on Underscore</td>
<td>No difference</td>
</tr>
</tbody>
</table>

---

1. D. Binkley, M. Davis, D. Lawrie, and C. Morrell, “To camelcase or underscore,” ICPC, 2009
Gender does NOT impact:
- Accuracy
- Time

- Spend more effort on alternative answers
- Seem to:
  - Carefully weight all options
  - Rule out wrong answers
- Statistically strong interaction between accuracy and effort on distracters

Seem to quickly go for one choice
- No correlation between effort and accuracy
Threats to Validity

- Internal validity:
  - Random ordering of source code
  - Comfortable environment

- External validity (generalization of the results):
  - Students as subjects
    - Students are the next generation of software professionals\(^1\)
  - 24 subjects
    - 9 female subjects → Not large enough to be generalized
    - A call for further studies

---
Gender does NOT impact:

- Accuracy
- Time

Male and female subjects use different strategies to select the correct answer.

Understanding systematic biases between men and women:

- Design tools and methods better adapted to different developers
- Support different program understanding strategies