Plagiarism in programming: Experts' and novices' beliefs about what it is and whether it is wrong

Katherine Hsiao[1,2], Ana McTaggart[2], Talia Waltzer[1], & Audun Dahl[1]

[1]Psychology Department, [2]Baskin School of Engineering

Project Overview

This project explores the relations between experience and perceptions of academic integrity in computer programming.

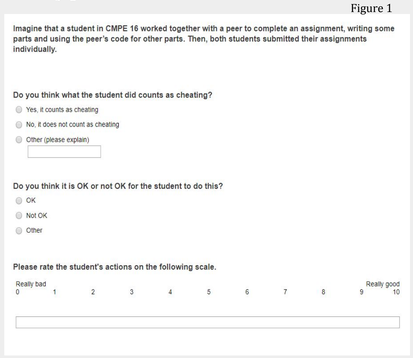
Demographics
- 12 Experts - *over 1 year of experience*
(18% Female, Mean age = 26.30)
- 19 Novices - *no coding experience*
(68% Female, Mean age = 21.80)

Procedure
1. 30-minute online survey

2. Is cheating wrong?

3. Hypothetical scenarios

Figure 1



4. Java Pair Programs

5. Ran the Java programs through MOSS

Figure 2

```java
/*
The following is the expected output:
"Please input five numbers"
"The sum of your list is __"
*/

import java.util.Scanner;

public class SeenByStudent {

  public static void main(String[] args) {
    int[] arr = new int[5];
    System.out.println("Please input five numbers");
    Scanner input = new Scanner(System.in);
    for (int i = 0; i < 5; i++) {
      int inp = input.nextInt();
      arr[i] = inp;
    }
//    for (int i = 0; i < 5; i++) {
//      System.out.println(arr[i]);
//    }
    int sum = 0;
    for (int i = 0; i < arr.length; i++) {
      sum += arr[i];
    }
    System.out.println("The sum of your list is " +sum);
  }
}
```

```java
/*
The following is the expected output:
"Please input five numbers"
"The sum of your list is __"
*/

import java.util.Scanner;

public class SubmittedByStudent {

  public static void main(String[] args) {
    int[] array = new int[5];
    System.out.println("Please input five numbers");
    Scanner sc = new Scanner(System.in);
    for (int i = 0; i < 5; i++) {
      int user = sc.nextInt();
      array[i] = user;
    }

    int total = 0;
    for (int i = 0; i < array.length; i++) {
      total += array[i];
    }

    System.out.println("The sum of your list is " +total);

  }
}
```

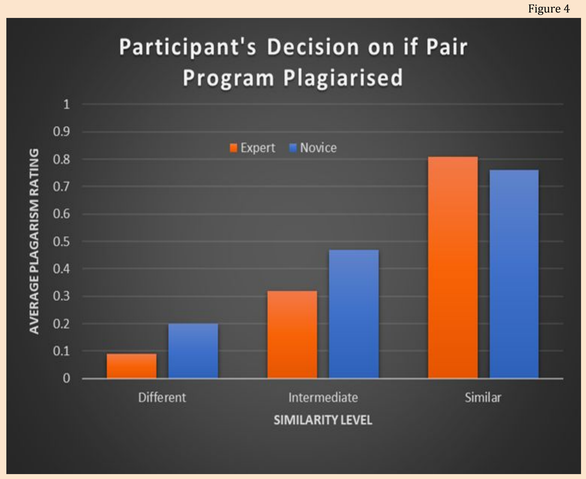
(1) What do individuals with varying programming experience believe counts as cheating?

Figure 3

Scenario type	Exact text of the hypothetical scenarios	% novices thought cheating	% experts thought cheating
OtherCoder	Imagine that a student in [CLASS] had someone else write a part of the program for them, and then turned it in as an assignment	95%	83%
Collaboration	Imagine that a student in [CLASS] worked together with a peer to complete an assignment, writing some parts and using the peer's code for other parts. Then, both students submitted their assignments individually	37%	70%
Debugging	Imagine that a student in [CLASS] went through a peer's code to fix it because it was not working	21%	27%
Website	Imagine that a student in [CLASS] looked on a website (such as StackOverflow) for a solution to a problem they didn't understand, then attempted to write the program and turned it in.	21%	10%
TestCode	Imagine that a student in [CLASS] used test code written by the professor to check whether an assignment was working properly	16%	0%

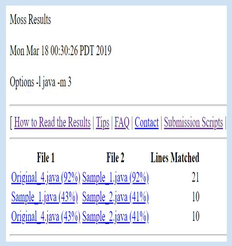
The differences highlight the importance of discipline specific experiences in classifying academic integrity.

Figure 4



Experts more clearly differentiated between all levels, and there was disagreement about intermediate cases.

(2) Do people's decisions align with those of a plagiarism detector (MOSS)?

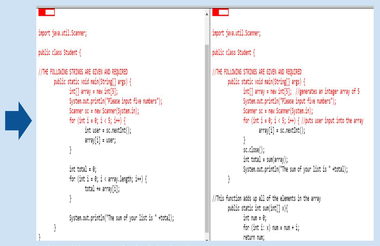

Moss Results

Mon Mar 18 00:30:26 PDT 2019

Options -l java -m 3

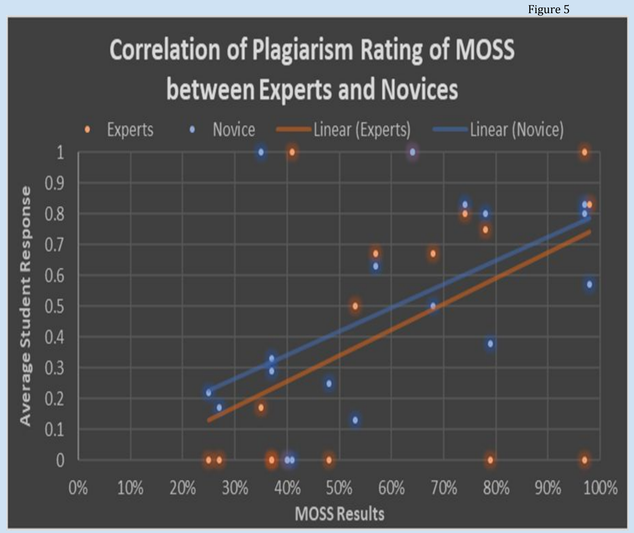
[How to Read the Results | Tips | FAQ | Contact | Submission Scripts]

File 1	File 2	Lines Matched
Original_4.java (92%)	Sample_1.java (92%)	21
Sample_1.java (43%)	Sample_2.java (41%)	10
Original_4.java (43%)	Sample_2.java (41%)	10



https://theory.stanford.edu/~aiken/moss/

Figure 5



*Expert: $r = .65$, $p = .004$ *Novice: $r = .56$, $p = .016$

While both groups positively correlated with MOSS, the plagiarism detector was slightly more representative of those with experience in programming than the novices.

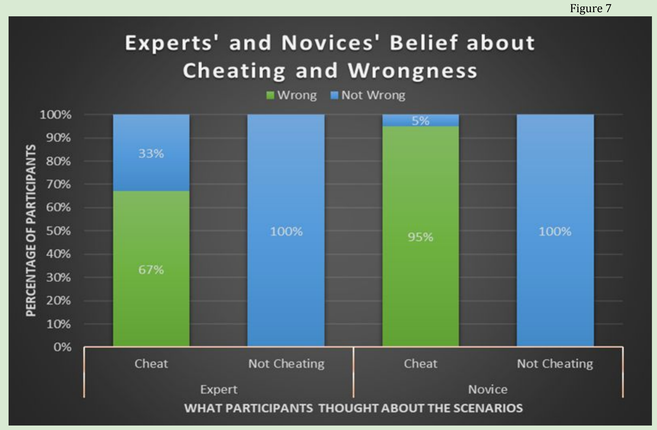
(3) Do people believe cheating is wrong?

Figure 6

"Cheating will misguide them and give them the illusion that learning is not important"

A majority of the participants believe that cheating is wrong, although some are more reserved about that claim.

Figure 7



*The Fisher exact test statistic value $p < .001$.

Compared to novices, experts had a higher percentage of circumstances in which acts of cheating would be okay

"If both understood the code and would be able to explain it fully, then it would be OK, in my opinion"

Discussion

- This method of examining plagiarism in programming has merit
- MOSS aligns closer with the expert's response
- There is a dichotomy between actions of cheating and whether it's okay in experts' perspective
- Indicates a different mindset towards plagiarism in those with engineering experience.

Next Steps

- Replicate this study with bracketed "Expert" group: Beginner, Intermediate, and Advanced programmers
- Explore cheating predictors by utilizing MOSS to analyze student programming assignments.

Acknowledgments

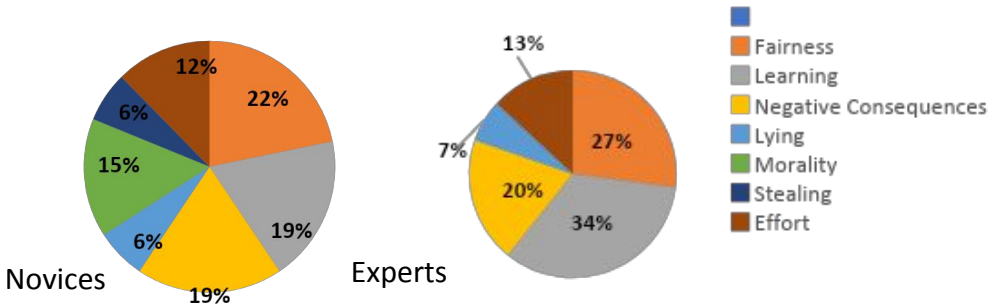
We would like to thank Max Wechsler-Azen, Arvid Samuelson, Maddy Dolinh, and Carmelle Bareket-Shavit. We also thank the Center for Innovations in Teaching and Learning, Baskin School of Engineering, and Psychology Department at UC Santa Cruz.

Correspondence: khsiao@ucsc.edu, twaltzer@ucsc.edu

Who were our participants?

Experts (*N* = 11)		Novices (*N* = 19)
Most experts were familiar with Red-Black Trees and Hashtables. Every expert knew conditional statements.	55% US Born	84% US Born / 11% not / 5% NA
	27% First-Gen College	26% First-Gen / 63% not / 11% NA
	45% Native English Speakers	79% Native / 16% not / 5% NA

Reasons given when asking whether cheating is wrong



Survey Questions

- Do you think what the student did counts as cheating?
- Do you think it is OK or not OK for the student to do this?
  - [If yes] Why is it OK?
  - [If no] Why isn't it OK?
  - [If other] Please explain your reasoning
- Please rate the student's action on the following scale

- Do you think the student's submission counts as plagiarism? [Image 1]
- How much do you think it is plagiarized?

"Second opinion's are just that. Usually you can give helpful hints once you spot their error."

"Sometimes getting code off the ground is hard, so looking at other codes helps you figure out what is need from that specific program."

"This is equivalent to the student copying old code from a Github repository, which I consider acceptable if they attempted to use other resources (office hours, talking with professor, reading textbook) first."

"I think the ability to solve problems using web-based resources is a fundamental skill"

```java
/*
The following is the expected output:
"Please input five numbers"
"The sum of your list is __"
*/

import java.util.Scanner;

public class SeenByStudent {

  public static void main(String[] args) {
    int[] arr = new int[5];
    System.out.println("Please input five numbers");
    Scanner input = new Scanner(System.in);
    for (int i = 0; i < 5; i++) {
      int inp = input.nextInt();
      arr[i] = inp;
    }
//    for (int i = 0; i < 5; i++) {
//      System.out.println(arr[i]);
//    }
    int sum = 0;
    for (int i = 0; i < arr.length; i++) {
      sum += arr[i];
    }
    System.out.println("The sum of your list is " +sum);
  }
}
```

```java
/*
The following is the expected output:
"Please input five numbers"
"The sum of your list is __"
*/

import java.util.Scanner;

public class SubmittedByStudent {

  public static void main(String[] args) {
    int[] array = new int[5];
    System.out.println("Please input five numbers");
    Scanner sc = new Scanner(System.in);
    for (int i = 0; i < 5; i++) {
      int user = sc.nextInt();
      array[i] = user;
    }

    int total = 0;
    for (int i = 0; i < array.length; i++) {
      total += array[i];
    }

    System.out.println("The sum of your list is " +total);

  }
}
```

[Image 1] This pair program is an example of a similar pair, where the variable name was change and comments were added.

Similar	Intermediate	Different
- Variable name change - Change some ordering of lines - Different comments	- May include Similar changes -Has 2-3 major changes, such as for loop to while loop, adding subfunctions, and/or extraneous code	- May include Intermediate changes -Has 4+ changes such as using different logic to approach the program, using different syntax, and/or different bracketing style

References

Parameswaran, A., & Devi, P. (2006). Student plagiarism and faculty responsibility in undergraduate engineering labs. *Higher Education Research & Development*, *25*(3), 263-276.

Schleimer, S., Wilkerson, D. S., & Aiken, A. (2003, June). Winnowing: local algorithms for document fingerprinting. In *Proceedings of the 2003 ACM SIGMOD international conference on Management of data* (pp. 76-85). ACM.

Waltzer, T., Hari, K., Gonzalez, G., Berman, E., & Dahl, A. (2017, April). When and why is it wrong to copy from others? Variability in students' evaluations of plagiarism. *Society for Research in Child Development Biennial Meeting,* Austin, TX.

Yeo, S. (2007). First-year university science and engineering students' understanding of plagiarism. *High Education Research & Development*, *26*(2), 199-216.

Correspondence: khsiao@ucsc.edu, twaltzer@ucsc.edu