Summary of Findings - Spring 2016 COP 3502 Cheating Survey

I asked the class about their perception of how often their peers cheated on programming assignments. (Full survey: CheatingSurvey.docx) Here is a summary of what I found:

128 responses, 121 answered more than the first question. I just used these for my analysis.

First of all, there's a huge standard deviation (21.2) in the students' estimates of what percentage of the class cheats. In fact, several students answered 0 and others answered 100, so the range was as large as possible in students answers. I assume those who answered 0 don't cheat and those who answered 100 do.

Next of all, there were some inconsistent answers. Namely, quite a few students answered the question, "What percentage of students in the class do you believe cheated?" with a lower percentage than the question: "What percentage of students in the class do you believe cheated in this manner(fill in)?" This is logically impossible since p(A) >= p(A and B). What I did for my survey data was that I screened out all students who made this error and multiplied their responses for questions of the second type by the answer of their first question. From there, I found the following averages for the class, in terms of what they perceive their peers to be doing:

30.2% cheat on programming assignments
11.1% cheat by getting a friend/acquaintance to write significant parts of the programs.
16.4% cheat by finding a copy of a solution online and modifying it slightly.
14.2% cheat by getting a copy of a current student's solution and modifying it slightly.
3.7% cheat by paying a person (possibly online) to write a program.

So, basically what we see is if we can trust crowd sourcing, about one third of students cheat on programming assignments and the most common way they cheat is by getting a copy that someone else has written (either an official solution from a past semester, another student's from a past semester, or another student from the current semester) and slightly editing the document they took.

Student Comments and Analysis

Most students believe that cheating brings up the class average. Many students specifically commented that this may artificially raise grades.

Many student brought up that it's unfair for one student to work many hours and get the same or worse grade than a cheating student.

One astute student actually considered the possibility that the cheaters learn less by cheating, then get lower grades on the exams that count more, which actually makes the class "curve" more favorable for the non-cheaters. An interesting perspective.

Many students pointed out that you can't cheat on the Foundation Exam so those students will either get stuck there, or learn the material somehow.

Also, many said students cheat because the assignments are too hard.

Student Quotes

"Cheating will affect me because I'll have to work with less-capable people because college failed to weed them out."

"Cheating is fine as long as you're learning from it."

"It hurts those who take the time to do the assignment. However, it the real world we can ask for help and google anything we want so we should always be able to do this."

"Class is too hard and lectures don't prepare for the assignments so we have to work together."

"Only cheating that could affect me is someone pays someone."

"Maybe some people who cheat don't think it's cheating."

"Collaboration and helping should be encouraged but not copy and paste."

"I get pressured to help others - if I did my help videos then all eyes would have been on me so I didn't do them, I don't need that while I am taking my own courses."

"Sean Szumlanski creates or moderates the Facebook page for his class which is where most people set times and destinations to collaborate and cheat."

"Coding I'd say is one of those things out there that you can actually learn from cheating because a perfect solution is almost always not out there. So, you have to fix it yourself indirectly learning it in the process."

"I haven't heard of others cheating in this course."

"It makes the effort and time put in to my program pointless. But then again, I get the advantage of learning"

"If the cheaters get caught and get 0s, it brings up my grade."

"I can't talk about programs with people who don't understand their own code."

"Students who cheat on programming assignments won't retain information as well as students who do the programs themselves. They won't succeed as well on the tests therefore, in a way, students who cheat improve my chances of succeeding in the course based on the grading system in place."

"Some of us don't have time to spend 5 hours to fix a seg fault."

"Not surprising, when we are told all that matters are grades."

"Cheapens the value of the degree."

"I think it's hard to complete a program without help."

My Analysis

I think students incorrectly assume that cheating helps those who engage in it. Ironically, this is probably what drives the rate of cheating up. Most wouldn't do it if they thought it wasn't going to help their grades, but obviously the reason they cheat is precisely because they think it will help their grades. In nearly all cases that I've caught students cheating on programs, they would have gotten roughly the same grade on that program had they worked on it on their own. Now, it may be the case that I disproportionately catch cheaters whose programs don't work, since those are the easier ones to catch. (The most common way we catch cheating is when two students do the same wrong weird identical thing…)

My change for Spring 2017 was to use some old programs but lower how much they counted for toward the course grade (I brought it all the way down to 2%, which is miniscule.) I added two quizzes on those programs worth 5% each, so answering analytical questions about the solutions to those programs was worth as much as doing all 5 of them. In this manner, I believe I've mitigated the effects (if there were any to begin with) that cheating has on average grades. My hope is that if the survey results are accurate, then 70% of students are still doing the work, then, they study my solution for their quiz, and they are really learning a fair bit about how to solve these canonical problems.

One other change I could provide is more stated opportunities to collaborate beyond the recitation programs. Unfortunately, students blew those off. I have to find a new trick so that they don't.

It's interesting that a few students thought it was cheating to use my posted code (for the class) or to talk about general high level ways to solve the problem. I actually tell them ***PRECISELY*** how to solve the problem in my hints and we still manage to get program averages in between 50% and 60% when 40% of the points pretty much tend to be for free. What's really fascinating is that students argue with me about how to solve the problems when I am the one who made them up. So, even when they are told how to solve something, they resist and sometimes don't follow the algorithm I ask them to.

Now, let's take a look at the five programs I assigned in Fall 2015, then again in Spring 2016 and then reassigned in Spring 2017. Here are the average grades on those programs:

Program	Fall 15 Avg	Spr 16 Avg	Spr 17 Avg
Word Search	50.5%	59.7%	51.2%
Mastermind	56.0%	55.2%	55.7%
Maze	60.8%	68.6%	54.9%
Boggle	57.1%	59.2%	65.1%

This data indicates that if cheating affected program averages significantly, then it only did so for the word search and maze programs in Spring of 2016. Though the Spring 2017 Boggle assignment looks like an outlier, the reason this average is different is because I provided a scaffold with a working trie and a fully perfect main for my students and they just needed to fill in the recursive method that solves the problem. In Fall 15 and Spring 16, there's no doubt that some students lost points for either (a) reading in the input incorrectly (forgetting to allocate space for the null

character in their grid) or (b) incorrectly implementing their trie. I have no doubt that students from those groups easily lost 6% to 8% on average, just due to those two reasons.

If anything, what this indicates is that one probably shouldn't reuse assignments in successive semesters, but anything beyond that doesn't seem to have enough of an effect. Even when using them in consecutive semesters, the effect doesn't show up on all assignments and is relatively modest.