ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE
Paper ID #36837

Cross-Sectional Survey of CS Students' Knowledge of and Attitudes Toward Cybersecurity

Cheryl Lynn Resch (Lecturer)

Cheryl Resch is an Instructional Assistant Professor in the Engineering Education Department at the University of Florida. She teaches core Computer Science courses and Cybersecurity courses in the Computer and Information Science and Engineering Department. Ms. Resch is also a PhD student in Human Centered Computing. Ms. Resch joined University of Florida in 2017. Prior to that she spent 29 years as an engineer at the Johns Hopkins University Applied Physics Laboratory. The last 15 years of her time at APL she worked on a wide variety of cybersecurity projects. Ms. Resch has a BS and MS in Mechanical Engineering from University of Maryland and an MS in Computer Science from Johns Hopkins University.

Christina Gardner-McCune

Keyna Wintjen

Cross-Sectional Survey of CS Students' Knowledge of and Attitudes Toward Cybersecurity

Abstract

Cyber attacks are a common feature of current news and many of them are the result of easy to avoid vulnerabilities in software. It is imperative that students graduating from an undergraduate Computer Science (CS) curriculum have knowledge of vulnerabilities and understand the consequences of vulnerable code. The introductory programming courses often have a full schedule of topics, so it is common to cover software security content briefly in early courses and more extensively in later courses. It would be useful to know if students with only a brief introduction to software vulnerabilities carry that knowledge into later classes. Also, it would be useful to have a sense of students' prior knowledge of cybersecurity and how this prior knowledge contributes to the appreciation of software vulnerabilities. This paper describes an analysis of the results of a survey of 1677 students in core CS courses at our large public university, in which software security topics are covered briefly in the two introductory courses and more extensively in a later course. We found that students in upper-level classes scored higher on a cybersecurity quiz than students who are just beginning the Computer Science curriculum, when correcting for prior knowledge or interest in cybersecurity. This suggests that students are gaining knowledge of cybersecurity while in the Computer Science curriculum. However, we found some gaps in cybersecurity knowledge. Less than half the students were able to name a software vulnerability that has caused a cybersecurity breach, and less than half were able to correctly answer questions about botnets and the use of VPNs. This suggests that we should consider increasing cybersecurity content in order to build on what students have apparently learned in the lower-level classes, and ensure that students learn all cybersecurity topics, and particularly about networking topics and common software vulnerabilities.

1 Introduction

Software vulnerabilities in commercial products are an issue of national security, financial and economic stability, and consumer confidence. Data breaches caused by these vulnerabilities can lead to interruptions in public services, monetary loss, and loss of privacy. The 2020 Verizon Data Breach Investigation Report [1] indicates that there were 3,950 data breaches in 2020 in the United States. Software vulnerabilities continue to increase as tracked by National Institute of Standards and Technology (NIST) National Vulnerability Database [2] and MITRE Common Vulnerabilities and Exposures (CVEs) [3]. A 2020 report from Tenable, a leading IT vulnerability assessment and management solution company, indicates that there were 18,358 vulnerabilities reported in 2020, a slight increase from the 17,305 reported in 2019 [4]. Despite increased tracking and abatement of software vulnerabilities, Gueye and Mell [5] report that the most prevalent software errors have not changed much since vulnerabilities were first cataloged. Indeed, MITRE [6] lists the top three software vulnerabilities as:

1. Improper Neutralization of Input on Webpage Generation (cross-site scripting)
2. Out-of-bounds write (buffer overflow)
3. Improper input validation

Software vulnerabilities can be reduced or eliminated when developers use principles of secure programming. It is vital that future developers are taught principles of cybersecurity and secure programming. The Association for Computing Machinery (ACM) included the Information Assurance and Security knowledge area in its Computer Science curriculum starting in 2008 [7].

The 2008 Computer Science curriculum included Foundational Concepts in Security (e.g., security goals of confidentiality, integrity and availability) Secure Programming, and Operating System Security. The 2013 Computer Science curriculum updated the knowledge area to also include Principles of Secure Design, Defensive Programming, Web Security, and Secure Software Engineering [8][9].

When determining how to distribute security topics in the Computer Science curriculum, it would be useful to have a sense of students' knowledge of and attitudes toward cybersecurity when they enter the curriculum, and how it changes as they progress through the curriculum. In this paper we explore the following research questions:

- RQ1 - Does students' knowledge of cybersecurity and software security correlate with where they are in the curriculum?
- RQ2 - Does students' attitude toward cybersecurity and software security correlate with where they are in the curriculum?

2 Prior Work

Our survey and analysis will draw on the work of Olmstead and Smith [10] who analyzed the results of a survey of the cybersecurity knowledge of the general public, and the work of Assal and Chiasson [11] who analyzed the results of a survey of attitudes of software developers. We will extend the work of Olmstead and Smith [10] by surveying Computer Science students, specifically, and by correlating the quiz results with how far they have proceeded through the curriculum, and whether they have a prior interest in cybersecurity. We will extend the work of Assal and Chiasson [11] by surveying Computer Science students rather than software developers. Previous literature has surveyed college students on their knowledge and attitudes towards cybersecurity [12][13][14][15][16][17]. The work described in this paper can be distinguished from that of this previous literature because it surveys Computer Science students specifically, and correlates the knowledge with prior interest in cybersecurity with which class they are taking.

3 Study Context

In this paper we analyze the responses of 1677 undergraduate students in six core Computer Science courses at a large R1 university to a survey of cybersecurity knowledge and attitudes. The six courses are: Programming Fundamentals 1 (CS1), Programming Fundamentals 2 (CS2), Advanced Programming Fundamentals (CS12), Computer Organization (CompOrg), Software Engineering (SoftEng), and Operating Systems (OS). The curriculum is designed such that students take one of these classes in each of their first five semesters. Currently, security topics are covered in modules of one to three lectures in Computer Organization, Software Engineering and Operating Systems classes.

The survey was optional and students received extra credit for participating. The study has been approved by the Institutional Review Board (IRB) at our university. Of the 1783 survey responses collected, 1677 were included in our analysis. Responses with missing demographic data were not included. Table 1 provides details on the courses from which participants were recruited and the number of participants from each course.

Table 1
Courses from which students were recruited

Abbreviation	Course Name	Brief Description	Number of Participants	Number of Students in Course	Percent Participation
CS1	Programming Fundamentals 1	First Course in Computer Science	597	725	82.3%
CS2	Programming Fundamentals 2	Second Course in Computer Science	313	347	90.2%
CS12	Advanced Programming Fundamentals	Once semester course covering CS1/CS2 for students with prior programming experience. Students must test in.	35	79	44.3%
CompOrg	Computer Organization	In third semester of curriculum. Software security covered briefly in this course.	384	440	87.2%
SoftEng	Software Engineering	In fourth semester of curriculum. Software security is covered in this class.	98	251	39.0%
OS	Operating Systems	In fifth semester of curriculum. Foundations of information security covered in this class.	250	347	72.0%
Total			1677	2189	

Participation was comparatively low for Software Engineering and Advanced Programming Fundamentals. It was discovered that many Software Engineering students were also taking Computer Organization or Operating Systems, and chose to receive the extra credit in those classes rather than Software Engineering. It is not clear why participation was so low in the Advanced Programming Fundamentals class. The author went to the class to promote the survey, the same as in the other classes. These students appear to not be as motivated to earn extra credit in the class.

To account for prior interest in cybersecurity, we asked participants to answer the following yes/no questions:

- Have you taken a cybersecurity class?
- Do you like to read about cybersecurity topics?
- Do you like to watch videos on cybersecurity topics?

Table 2 indicates how many students answered "yes" to each of these questions.

Table 2
Number of Students who Responded "Yes" to Questions about Cybersecurity Interest

Interest in Cybersecurity	Number of Participants	Percentage of participants
Have taken a cybersecurity class	116 out of 1677	6.9%
Read about cybersecurity	717 out of 1677	42.8%
Watch videos on cybersecurity	800 out of 1677	47.7%

3.1 Survey contents and results

To measure participants' knowledge of general cybersecurity topics, we used questions from a Pew Research survey meant for the general public [10]. The Pew Research Survey [10] included the questions shown in Table 3. This survey collected responses from 1055 adult internet users in 2016. We supplemented these questions with the following:

- Name a common software vulnerability that has resulted in a security breach.

A correct answer was given a score of 1 and an incorrect answer was given a score of 0. For the question that asked participants to name a software vulnerability, the answer was scored as a 1 if the answer could be interpreted as a software vulnerability. Examples of correct answers to the software vulnerability question are input validation, cross site scripting, code injection, improper authentication, and weak authentication allowed. Answers were deemed incorrect if they did not pertain to software. Some examples of incorrect answers are phishing and using the same password for multiple websites. Blank answers and answers indicating they did not know were also scored 0. The mean score is the fraction of students' responses that were correct. The questions about two-factor authentication, secure passwords, ransomware, and safety of public WiFi had the highest fraction of correct responses. The questions asking about what risks VPNs mitigate, and the open-ended question asking students to name a software vulnerability had the lowest fraction of correct answers. The last column indicates the results from a survey of 1055 adult internet users in 2016. The fraction of correct answers from this survey of the general public are all lower than those of our survey of Computer Science students. The trends are similar, with VPNs and botnets getting lower scores. In 2016, the survey of the general public showed only 10% could identify 2-factor authentication, while 89% of the students surveyed could identify 2-factor authentication. Our university requires 2-factor authentication for access to the learning management system and email, which may explain the high percentage of survey respondents who could identify it.

Table 3

Mean Scores for Survey Responses

Question	Abbrev	Choices (correct answer in bold)	Mean Score	Results from [10]
What does the “https://” at the beginning of a URL denote, as opposed to "http://" (without the “s”)?	https	● This is the newest version available ● This site is not accessible to certain computers ● **This site is encrypted** ● This site has special high definition ● All of the above ● Not sure	0.659	0.33
A group of computers that is networked together and used by hackers to steal information is called	botnet	● Operating system ● **Botnet** ● DDoS ● Not sure	0.585	0.16
Some websites and online services use a security process called two-factor authentication. Which of the following images is an example of two-factor authentication?	twofactor		0.893	0.10
Which of the following four passwords is the most secure?	password		0.962	0.75
Criminals access someone's computer and encrypt the user's personal files and data. The user is unable to access this data unless they pay the criminals to decrypt the files. This practice is called....	ransom	● Driving ● Spam ● **Ransomware** ● Botnet ● None of these ● Not sure	0.842	0.48
If a public Wi-Fi network (such as in an airport or cafe) requires a password to access, is it generally safe to use that network for sensitive activities such as online banking?	wifi	● **No, it is not safe** ● Yes, it is safe ● Not sure	0.859	0.73
What kind of cybersecurity risks can be minimized by using a Virtual Private Network (VPN)?	vpn	● Keylogging ● Phishing ● De-anonymization by network operators ● **Use of insecure Wifi networks** ● Not sure	0.472	0.13
Name a common software vulnerability that has resulted in a security breach	softvuln		0.330	
total			5.526	

To gauge participants' attitudes about cybersecurity topics in the curriculum, we asked participants to answer whether “All Computer Science students should learn” each of the topics shown in Table 4, with a 5-point Likert scale. For each question there was also a 6th option in which students could answer “Not familiar with this concept” that was scored as 0. The list of topics is from the ACM Curriculum Information Assurance and Security knowledge area. For participants who gave an answer to all the questions, we summed their responses to obtain one measure of each student’s attitude about the necessity of learning cybersecurity.

Table 4
Mean Value of Responses for Topics Computer Science Students Should Study

Computer Science Students Should Learn About	**Observations**	**Mean**
Common vulnerabilities	1635	4.600
Adversarial thinking	1629	3.746
Penetration testing	1632	3.920
Security policies	1631	4.368
Authentication and access control	1634	4.328
Formal methods	1632	3.707
Algorithm soundness and completeness	1635	4.284
Cryptography	1634	3.963
SumShould	**1610**	**32.932**

The high values indicate that most students strongly agree that students should learn about each of the cybersecurity topics.

We also asked participants to rate their own knowledge of the cybersecurity topics shown in Table 5 with a 5-point Likert scale. For each question there was also a 6th option in which students could answer “Not familiar with this concept” that was scored as a 0. For participants who gave an answer to all the questions, we summed their responses to obtain one measure of each student’s self-assessment of cybersecurity knowledge.

Table 5
Mean Value of Responses for Topics Participants are Knowledgeable About

I Am Knowledgeable About	Observations	Mean
Common vulnerabilities	1637	2.429
Adversarial thinking	1633	1.805
Penetration testing	1630	1.812
Security policies	1625	2.140
Authentication and access control	1635	2.259
Formal methods	1633	1.656
Algorithm soundness and completeness	1633	2.203
Cryptography	1633	2.202
SumKnow	**1600**	**16.325**

These values are quite low, indicating that the average response for all the topics was that participants do not think they are knowledgeable about these topics.

3.2 Model and Analysis

We seek to answer the following research questions:

- Does students' knowledge of cybersecurity and software vulnerabilities correlate with where they are in the curriculum? We explore this question with three hypotheses:
    - H1a - Do students in CS2, CS12, CompOrg, SoftEng, and OS have more knowledge of cybersecurity than students in CS1?
    - H1b - Is the cybersecurity knowledge of students in CompOrg different from students in OS?
    - H1c - Do students in CompOrg, SoftEng and OS have more knowledge of cybersecurity than students in CS1, CS2, and CS12?
- Does students' attitude toward cybersecurity correlate with where they are in the curriculum? We explore this question with three hypotheses:
    - H2a - Do students in CS2, CS12, CompOrg, SoftEng and OS have a different attitude toward cybersecurity than students in CS1?
    - H2b - Do students in CompOrg have a different attitude toward cybersecurity than students in OS?
    - H2c - Do students in CompOrg, SoftEng and OS have a different attitude toward cybersecurity than students in CS1, CS2, and CS12?

To answer H1a, H1b, and H1c, we summed each participant's scores on the cybersecurity quiz questions found in Table 3 and called that variable *total*. To answer H1a and H1b, we employed a multiple linear regression model that predicted *total* as a function of which course they are taking, while correcting for whether students had taken a class on cybersecurity (variable *Class*), like to read about cybersecurity (variable *Read)*, or like to watch videos on cybersecurity

(variable *Video*). CS1 serves as the base variable since it is the first course in the CS curriculum and assumes no prior knowledge.

$$total = \beta_0 + \beta_1 CS2 + \beta_2 CS12 + \beta_3 CompOrg + \beta_4 SoftEng + \beta_5 OS + \beta_6 Class + \beta_7 Read + \beta_8 Video + u \quad \text{(Equation 1)}$$

To answer H1c, we created a binary variable, *upper,* that is coded as one if the participant was currently enrolled in CompOrg, SoftEng, or OS and zero otherwise. We employed the following regression model:

$$total = \beta_0 + \beta_1 upper + \beta_2 Class + \beta_3 Read + \beta_4 Video + u \quad \text{(Equation 2)}$$

To obtain more insight on performance on individual questions that participants performed poorly on, we employed logistic regression to predict the mean score of individual quiz questions.

$$meanscore = \beta_0 + \beta_1 upper + \beta_2 Class + \beta_3 Read + \beta_4 Video + u \quad \text{(Equation 3)}$$

We calculated regression coefficients for each quiz question.

To answer H2a and H2b, we employed a multiple regression model that predicted participants' self-assessment of their own cybersecurity knowledge, and their attitude about the necessity of learning cybersecurity, as a function of which course they are taking, while correcting for whether students had taken a class on cybersecurity, like to read about cybersecurity, or like to watch videos on cybersecurity. The variable *SumShould* is the sum of a participant's answers to the Likert questions in Table 4. The variable *SumKnow* is the sum of a participant's answers to the Likert questions in Table 5.

$$SumShould = \beta_0 + \beta_1 CS2 + \beta_2 CS12 + \beta_3 CompOrg + \beta_4 SoftEng + \beta_5 OS + \beta_6 Class + \beta_7 Read + \beta_8 Video + u \quad \text{(Equation 4)}$$

$$SumKnow = \beta_0 + \beta_1 CS2 + \beta_2 CS12 + \beta_3 CompOrg + \beta_4 SoftEng + \beta_5 OS + \beta_6 Class + \beta_7 Read + \beta_8 Video + u \quad \text{(Equation 5)}$$

To answer H2c, we used a multiple regression model to model *SumShould* and *SumKnow* as a function of the *upper* variable.

$$SumShould = \beta_0 + \beta_1 upper + \beta_2 Class + \beta_3 Read + \beta_4 Video + u \quad \text{(Equation 6)}$$

$$SumKnow = \beta_0 + \beta_1 upper + \beta_2 Class + \beta_3 Read + \beta_4 Video + u \quad \text{(Equation 7)}$$

3.3 Assumptions

Our population is computer science, computer engineering, and digital arts and science students at a large R1 university. Our sample is drawn from all students enrolled in six different required core courses offered in different semesters of the curriculum. The total population of students in these majors is 3055, and 2197 students had the survey made available to them. The sampling is not truly random because participants are self-selected and motivated to obtain extra credit. However, over half of the total population, and 76% of the population with the opportunity to participate did so. The number of participants and percentage of population that participated is high, so bias due to nonrandom sampling will be minimal. The Shapiro-Wilk W [14] test indicates that the data are not normally distributed. Based on this formal test of normality and

the resulting significant p-value, the data are not normally distributed, but due to the large sample size, we are comfortable in relaxing the normality assumption. White's [15] test indicates that homoscedasticity is not met. Because the assumption of homoscedasticity is not met, we will report robust standard error. Scatterplots indicate that *total, SumKnow*, and *SumShould* are linear in the parameter *coursenum*. Scatter plots for the residuals from the models indicate that the assumption of zero conditional mean is met.

4. Findings

Table 6 presents the multiple regression results for Equation 1, *total* versus a regression of *coursenum, Class, Read, and Video*.

Table 6
Multiple Regression Results for Equation 1

Variable	Coefficient	Robust Standard Error
constant	4.918*	0.068
Class	0.534*	0.125
Read	0.526*	0.071
Video	0.429*	0.071
CS2	0.193*	0.093
CS12	0.624*	0.226
CompOrg	0.348*	0.148
SoftEng	0.388*	0.148
OS	0.512*	0.100

* p<0.05

Class, Read, Video have a significant effect on total score on the quiz. Each factor adds about a half a point to the total score. Students in CS2, CS12, CompOrg, SoftEng and OS score significantly higher than students in CS1 on the cybersecurity quiz, after correcting for whether they have an interest in cybersecurity. The highest factors are for the students in CS12, Advanced Programing Fundamentals, and OS, Operating Systems.

To answer H1b, we ran a Wald test on the null hypothesis that the coefficient *CompOrg* and *OS* are the same. The result, $F(1, 1668) = 2.18,\ p = 0.1396$, indicates that we cannot reject the null hypothesis. There is no significant difference in the coefficient for *CompOrg* and *OS* in predicting the *total* variable.

Table 7 presents the multiple regression results for Equation 2, *total* versus a regression of *upper, Class, Read, and Video.*

Table 7
Multiple Regression Results for Equation 2

Variable	Coefficient	Robust Standard Error
constant	5.005*	0.059
Class	0.546*	0.125
Read	0.537*	0.071
Video	0.424*	0.072
upper	0.320*	0.067

* p<0.05

Class, Read, and *Video* have a significant effect on the total score. Being in a later course in the curriculum also has a significant effect on the total score on the quiz, when correcting for *Class, Read,* and *Video*.

Regression analysis was performed for each of the quiz questions using Equation 3. Tables of regression coefficients for each quiz question are found in the appendix. Table 8 summarizes the results. Students in CompOrg, SoftEng, and OS were significantly more likely than students in CS1, CS2, and CS12 to correctly answer questions about https, and wifi, and were significantly more likely to be able to name a software vulnerability. For the two questions with the lowest mean scores, on botnets and VPNs, being a student in CompOrg, SoftEng, and OS did not have a significant effect.

Table 8
Significant Factors on Regression for Individual Quiz Questions

Quiz Question	Mean Score	Significant Factors
https	0.659	upper, Class, Read, Video
Botnet	0.585	Class, Read, Video
twofactor	0.893	Read
password	0.962	
ransom	0.842	Class, Read, Video
Wifi	0.859	upper, Read, Video.
Vpn	0.472	Class
Softvuln	0.330	upper, Read, Video

Table 9 presents the multiple regression results for the variable *SumShould* as a function of *coursenum, Class, Read, and Video.*

Table 9
Multiple Regression Results for Equation 4

Variable	Coefficient	Robust Standard Error
constant	29.178*	0.555
Class	1.350*	0.545
Read	2.492*	0.428
Video	2.200*	0.451
CS2	-0.296	0.711
CS12	-0.341	1.858
CompOrg	0.691	0.596
SoftEng	-1.346	1.063
OS	0.600	0.671

$R^2 = 0.0458$, * $p<0.05$

The analysis shows that there is no significant effect on the variable *SumShould* for students in any class compared to students in CS1. To answer H2b, we ran a Wald test on the null hypothesis that the coefficient CompOrg and OS are the same. The result, $F(1, 1668) = 0.02$, $p = 0.8942$, indicates that we cannot reject the null hypothesis. There is no significant difference in the coefficient for CompOrg and OS in predicting the *SumShould* variable.

Table 10 presents the multiple regression results for the variable *SumKnow* as a function of *coursenum, Class, Read, and Video*

Table 10
Multiple Regression Results for Equation 5

Variable	Coefficient	Robust Standard Error
constant	9.718*	0.433
Class	7.425*	0.790
Read	4.713*	0.445
Video	4.087*	0.441
CS2	0.826	0.615
CS12	1.102	1.450
CompOrg	2.599*	0.544
SoftEng	2.968*	0.885
OS	3.459*	0.584

* $p<0.05$

The analysis shows that students in CompOrg, SoftEng, and OS rate their own knowledge of cybersecurity significantly higher than that of students in CS1 when controlling for variables *Class, Read, Video*.

To answer H2b, we ran a Wald test on the null hypothesis that the coefficient CompOrg and OS are the same. The result, $F(1, 1668) = 2.09$, $p = 0.1487$, indicates that we cannot reject the null hypothesis. There is no significant difference in the coefficient for CompOrg and OS in predicting the *SumKnow* variable.

Table 11 presents the multiple regression results for the variable should as a function of *upper, Class, Read, and Video*

Table 11
Multiple Regression Results for Equation 6

Variable	Coefficient	Robust Standard Error
constant	29.07	0.504
upper	0.499	0.455
class	1.34*	0.538
read	2.47*	0.425
video	2.21*	0.449

* $p<0.05$

The analysis shows that there is no significant effect on the variable *SumShould* for students in CompOrg, SoftEng, and OS compared to students in CS1, CS2, and CS12. Table 12 presents the multiple regression results for the variable *SumKnow* as a function of *upper*, *Class, Read, and Video*

Table 12
Multiple Regression Results for Equation 7

Variable	**Coefficient**	**Robust Standard Error**
constant	10.04*	0.371
upper	2.63*	0.403
class	7.52*	0.773
read	4.74*	0.443
video	4.04*	0.440

* $p<0.05$

The analysis shows that students in CompOrg, SoftEng,and OS rate their knowledge of cybersecurity significantly higher than that of students in CS1, CS2, CS12.

4.1 Summary of Results

To answer the question "Does students' knowledge of cybersecurity and software vulnerabilities correlate with where they are in the curriculum?", we found that being in CS2, CS12, CompOrg, SofttEng and OS had a significant effect on the total score on a cybersecurity quiz compared to students in CS1, when correcting for prior interest in cybersecurity. There was no significant difference in the score on the cybersecurity quiz for students in CompOrg compared to students in OS. Similarly, we found that there is a significant difference in scores on the cybersecurity quiz for students in CompOrg, SoftEng, and OS compared to students in CS1, CS2, and CS12, when correcting for prior interest in cybersecurity. Students in all core courses scored significantly better on the quiz than students in the first course, CS1. Interestingly, students in CS12, the advanced programming fundamentals course that students had to test into, had the highest coefficient. The next highest was for students in the OS class, which is the course latest in the curriculum. Students in the later courses in the curriculum were significantly more likely to be able to name a software vulnerability, but not significantly more likely to correctly answer questions about botnets or VPNs.

To answer the question "Does students' attitude toward cybersecurity correlate with where they are in the curriculum?", we found that students in CompOrg, SoftEng, and OS rate their knowledge of cybersecurity significantly higher than students in CS1. However, there was no significant difference in students' rating of the importance of learning cybersecurity students in any class compared to students in CS1. We found no significant difference in how students rate their own knowledge of cybersecurity for students in CompOrg compared to students in OS and no significant difference in the rating of importance of learning cybersecurity for students in CompOrg compared to students in OS. We found a significant difference in students' rating of their own cybersecurity knowledge for students in the later classes in the curriculum compared to

students in CS1, CS2, and CS12. There was no significant difference in students' rating of the importance of cybersecurity for students in the later courses in the curriculum compared to students in CS1, CS2, and CS12. We can conclude that students in later courses in the curriculum rate their own knowledge of cybersecurity higher than that of students earlier in the curriculum, but all students rate the importance of learning cybersecurity very high.

4.2 Limitations / Threats to Validity

Although the Pew Research Center is well-known for its methodology in developing and administering surveys world-wide, we did not find a measure of the internal consistency, i.e., Cronbach's alpha, of the Pew cybersecurity quiz we adopted, nor did we test the reliability of our survey after adding the additional questions. This limits our ability to generalize to other populations without a measure of reliability. Additionally, we found that our multiple linear regression model did not meet the normality assumption. However, due to the large sample size, we relaxed the assumption of normality. More research is needed to determine how demographics might affect student scores on the cybersecurity quiz, thus a disaggregation across demographics would be prudent. It would also be interesting to conduct a post test with students in the introductory computer science courses at the end of the semester to see whether scores on the cybersecurity quiz increase.

5 Conclusion

Our analysis found that students in upper level classes have a higher level of knowledge of cybersecurity than students who are just beginning the Computer Science curriculum, when correcting for a prior knowledge or interest in cybersecurity. This suggests that students are gaining knowledge of cybersecurity while in the Computer Science curriculum. Indeed the regression coefficients seem to indicate that cybersecurity knowledge is higher the further students are in the curriculum. However, there are some gaps in cybersecurity knowledge, as many students at all levels were not able to name a software vulnerability, and were not able to correctly answer questions about botnets and the use of VPN. We should consider increasing cybersecurity content in classes in the middle of the curriculum in order to build on what students have apparently learned in the lower level classes, and to ensure that students learn all cybersecurity topics, and particularly about common software vulnerabilities and how to prevent them. Our analysis showed that students in upper level courses rate their own knowledge of cybersecurity higher than students in the lower level classes when correcting for prior knowledge and interest in cybersecurity. We found no differences in the rating of importance of cybersecurity topics among students in the different classes, when correcting for prior knowledge and interest in cybersecurity. The average value for this parameter was quite high, indicating that students at all levels understand the importance of learning cybersecurity. In conclusion, our analysis showed that students in upper level courses in our curriculum scored higher on a cybersecurity quiz than students in introductory courses, and students in upper level courses rate their knowledge of cybersecurity higher than students in introductory courses. Students in all courses agree that cybersecurity topics are important to learn. These are encouraging results. Students appear to gain cybersecurity knowledge, and understand the importance of learning about cybersecurity. We plan to have students who are now in the introductory courses take the survey again to determine if their scores increase. We also plan to introduce course content early in the curriculum to address the apparent gaps in knowledge about software vulnerabilities and the use of VPNs.

References

[1] Bassett, G., Hylender, D., Langlois, P., Pinto, A., & Widup, S., "Data Breach Investigations Report. Retrieved" from:https://www.verizon.com/business/resources/reports/dbir, 2020.

[2] NIST National Vulnerability Database.. Retrieved from: https://nvd.nist.gov/vuln/data-feeds, Accessed February 1, 2022.

[3] MITRE Common Vulnerabilities and Exposures. Retrieved from: https://cve.mitre.org/cve/data_feeds.html, Accessed February 1, 2022.

[4] "Tenable's 2020 Threat Landscape Retrospective" Retrieved from: https://www.tenable.com/cyber-exposure/2020-threat-landscape-retrospective, Accessed February 1, 2022.

[5] Gueye, A., & Mell, P, "A Historical and Statistical Study of the Software Vulnerability Landscape," *arXiv preprint arXiv:2102.01722*, 2021.

[6] MITRE 2020 CWE Top 25 Most Dangerous Software Errors, 2021, Available: https://cwe.mitre.org/top25/archive/2020/2020_cwe_top25.html, Accessed February 1, 2022.

[7] Cooper, S., Nickell, C., Piotrowski, V., Oldfield, B., Abdallah, A., Bishop, M., & Brynielsson, J., "An exploration of the current state of information assurance education" *ACM SIGCSE Bulletin*, *41*(4), 109-125, 2010.

[8] ACM Curriculum guidelines for undergraduate degree programs in computer science 2013, Retrieved from:https://www.acm.org/binaries/content/assets/education/., Accessed February 1, 2022.

[9] Yuan, X., Yang, L., Jones, B., Yu, H., & Chu, B. T., "Secure software engineering education: Knowledge area, curriculum and resources," *Journal of Cybersecurity Education, Research and Practice*, *2016*(1), 3, 2016.

[10] Olmstead, K., & Smith, A., "What the public knows about cybersecurity," *Pew Research Center*, *22*, 2017.

[11] Assal, H., & Chiasson, S.,. "'Think secure from the beginning' A Survey with Software Developers," In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1-13), May 2019.

[12] Farooq, A., Isoaho, J., Virtanen, S., & Isoaho, J., Information security awareness in educational institution: An analysis of students' individual factors. In *2015 IEEE Trustcom/BigDataSE/ISPA*, Vol. 1, pp. 352-359, 2015.

[13] Kim, E. B., "Recommendations for information security awareness training for college students," *Information Management & Computer Security*, 2014.

[14] Farooq, A., Kakakhel, S. R. U., Virtanen, S., & Isoaho, J., "A taxonomy of perceived information security and privacy threats among IT security students," In *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, pp. 280-286, 2015.

[15] Bhatnagar, N., & Pry, M., "Student Attitudes, Awareness, and Perceptions of Personal Privacy and Cybersecurity in the Use of Social Media: An Initial Study," *Information Systems Education Journal*, *18*(1), 48-58, 2020.

[16] Chandarman, R., & Van Niekerk, B. "Students' cybersecurity awareness at a private tertiary educational institution," *The African Journal of Information and Communication*, *20*, 133-155, 2017,

[17] Gabra, A. A., Sirat, M. B., Hajar, S., & Dauda, I. B., "Cyber security awareness among university students: A case study," *Journal of Critical Reviews*, *7*, 16, 2020.

[18] Shapiro, S. S., & Wilk, M. B., :An analysis of variance test for normality (complete samples," *Biometrika*, 52, 591-611, 1965.

[19] White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* vol 48 no 4, pp 817–838, 1980.