ATTITUDES ON PLAGIARISM IN PROGRAMMING COURSES: RESULTS FROM A SURVEY OF STUDENT PERCEPTIONS

Cheryl Aasheim  
Georgia Southern University  
P.O. Box 8150  
Statesboro, GA 30460  
912-478-5165  
caasheim@georgiasouthern.edu

Paige Rutner  
Georgia Southern University  
P.O. Box 7998  
Statesboro, GA 30460  
912-478-5924  
prutner@georgiasouthern.edu

Lixin Li  
Georgia Southern University  
P.O. Box 7997  
Statesboro, GA 30460  
912-478-7646  
lli@georgiasouthern.edu

INTRODUCTION

There have been many studies conducted on plagiarism in academic institutions. Studies include investigation of plagiarism in different groups of students as well as using different methods to plagiarize. Sheard, Dick, Markham, et al. (2002) studied plagiarism of first year IT students while McCabe, Butterfield & Trevino (2006) investigated academic dishonesty in graduate business programs. Cosma & Joy and Daly & Horgan discussed detecting plagiarism of source code in programming classes. Given that many assignments in academic institutions are submitted electronically and the availability of information in digital format, plagiarism is, arguably, easier today than it has been in the past.

This research seeks to add to the current body of literature by investigating whether college students apply the same standards of acceptability for graded programming assignments versus other types of graded assignments. The rationale for this research objective is that the authors of this study are faculty who teach programming and have done so for years; anecdotally, we have noticed that cheating on programming assignments is common. This observation provided the initial impetus for this study. In addition, we have discovered that students do not typically think of code they produce like they would an essay. They seem to think copying a portion of code is not unethical while copying a portion of an essay is a clear violation of the university academic integrity policy. There have been a number of studies citing similar observations with respect to programming assignments and other computer based work (Ross 2005; Buchanan 2006; Cosma and Joy 2008; Marsan 2010). However, there have been few researchers that have compared cheating on IT based assignments to cheating on traditional assignments; furthermore, the findings from the few studies that have looked at this issue are not consistent. Molnar et al (2008) found that undergraduate students rated cheating with IT as more acceptable than cheating without the use of IT. Stephens et al. (2007), on the other hand, found that students did not express different perceptions about the seriousness of cheating in a digital context versus a standard context.
To accomplish the research goal, a survey was conducted on student perceptions of the acceptability of a number of behaviors when working on various types of graded individual assignments. The types of assignments investigated are: (1) programming, (2) mathematics, and (3) essay. The behaviors included in the survey are based upon the following categories previously identified by Sheard et al. (2002), Broeckelman-Post (2008) and Jian et al. (2008): (1) seeking help from approved sources, (2) unauthorized collaboration, (3) copying portions of others’ work, and (4) copying all of others’ work.

BACKGROUND AND LITERATURE REVIEW

Cheating and plagiarism by college students is an area of concern to academics both in their capacity as teachers and as researchers Academic research in this area has a long tradition with some of the earliest works dating back to the early years of the 20th century (e.g., Barnes 1904; Campbell 1933; Drake 1941). As might be expected, the body of literature on this topic is extensive and a full review is beyond the scope of the current paper. However, we present an overview of research on cheating with special attention to the work most relevant to the current study. We divide this literature review into two broad areas. In the first area, we group those articles that explore the prevalence of cheating and that attempt to determine the extent to which personal and environmental factors influence cheating. In the second area, we group articles assessing the impact of technology and the internet on cheating.

Studies exploring the prevalence of cheating have found wide ranging results. The percentage of students who admit to some form of academic dishonesty range from a low of 3% (Karlins, Michaels et al. 1988) to a high of 95% (McCabe and Trevino 1997). The disparities in cheating rates found in these studies can be attributed a variety of factors. They encompass different definitions of cheating and plagiarism, different methods of measurement, and different types of student work. While much of the work in this area depends on self-report measures to determine the rate of cheating (McCabe, Trevino et al. 2001), other studies using measures of actual cheating behavior have also identified a broad range in rates of cheating. One of the early studies attempting to determine actual cheating behavior found the low 3% rate mentioned previously (Karlins, Michaels et al. 1988). West, Ravenscroft, and Schrader (2004) used an incident of widespread, blatant cheating (74% of a class cheated on a take-home exam) to conduct a natural experiment comparing cheating to measures of moral judgment. The advent of widely available text matching software tools such as Turnitin has increased the number of studies reporting rates of actual cheating behavior detected through use of the tools: these studies have reported rates ranging from 21% to 61% (Warn 2006; Ledwith and Risquez 2008; Martin, Rao et al. 2009; Walker 2010). Despite the attention given academic dishonesty, the rate of occurrence does not appear to be declining and may be increasing (Haines, Diekhoff et al. 1986; Park 2003; Eastman, Iyer et al. 2006).

Technology has had a profound impact on the academic environment providing greater access to students in widespread locations and improving the ease of communicating and disseminating information (Mayfield and Ali 1996). However, technology has also increased the opportunity and ease of student cheating. The extant literature contains numerous examples of students using technology to gain easy access to other’s work or solicit unauthorized assistance. The most egregious forms of student cheating are the outright purchase of assignments such as term papers (Campbell, Swift et al. 2000) or completed assignments (Ross 2005). Other forms of cheating include copying and pasting unattributed material from online sources. Today’s college students have grown up using the Internet as a primary source of information (Marsan 2010); we should not be surprised when they turn to it when working on assignments.

Despite the recognition of how technology and the Internet have enabled increased cheating (Renard 1999/2000; Ercegovac and Richardson 2004), there has been relatively little academic literature offering
empirical examinations of the phenomenon. Lester and Diekhoff (2002) conducted one of the earliest studies comparing demographic and other factors of traditional and Internet cheaters. The study found that where traditional cheaters tended to be women (65.2%), Internet cheaters were more likely to be men (54.1%). Internet cheaters were also more likely to be involved with both varsity and intramural sports than traditional cheaters but no significant differences were found in other demographic factors. This study found that both traditional and Internet cheaters tended to justify their behaviors and that this justification was more prevalent in Internet cheaters. Finally, the study found that Internet cheaters were less likely to resent cheating behavior in others than were traditional cheaters.

**METHODOLOGY**

This study combines and builds on the results of two previous studies conducted by the authors (Aasheim, Li, Rutner and Williams, 2010; Aasheim, Li and Rutner, 2011). The results of the first study (Aasheim et al., 2010) were based on a preliminary investigation of data and indicated that education about academically dishonest behavior related to programming did seem to make a difference on student perceptions. In the second study (Aasheim et al., 2011) the authors conducted a factor analysis to determine if the underlying dimensions for the three types of assignments were different, which would provide initial support for the hypothesis that students do view the types of assignments differently. The results indicated that students did seem to have different perceptions about how they view ethical behaviors as they relate to the different types of assignments. This study is to develop and test hypotheses related to how students view the different types of assignments differently. To this end the following hypothesis is developed and tested:

Students have different perceptions on what constitutes academically dishonest behavior depending on the type of graded assignment (programming, math or English).

As programming assignments are the primary concern, the authors wish to determine whether perceptions on programming assignments differ from those of essay or math assignments.

\[H1: \text{Students have different perceptions on what constitutes academically dishonest behavior for programming assignments than math assignments.}\]

\[H2: \text{Students have different perceptions on what constitutes academically dishonest behavior for programming assignments than English assignments.}\]

After examining current literature (Sheard, Markham et al. 2002; Broeckelman-Post 2008; Jian, Sandnes et al. 2008) on categories of academic behavior related to graded class assignments, three sets of 12 questions for programming, mathematical and essay assignments were designed in the survey (see Appendix 1). The mapping of the categories according to literature and the questions on the survey is provided in Table 1.

<table>
<thead>
<tr>
<th>Categories of Academic Behavior</th>
<th>Survey Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeking help from approved sources</td>
<td>1 - 3</td>
</tr>
<tr>
<td>Unauthorized collaboration</td>
<td>4 - 6</td>
</tr>
<tr>
<td>Copying portions of others’ work</td>
<td>7 - 9</td>
</tr>
<tr>
<td>Copying all of others’ work</td>
<td>10 - 12</td>
</tr>
</tbody>
</table>
The survey instrument was administered to undergraduate students in programming courses at the authors’ institution. For a graded programming, math and essay assignment that were to be completed individually, the respondents were asked to indicate (on a Likert scale of 1=Very Acceptable Behavior to 5=Very Unacceptable Behavior) how acceptable they felt the behaviors specified in the survey questions were. The survey was a retrospective pre-test/post-test instrument in that it was given two points in time: (1) at the beginning of enrolling in the programming course and (2) after class discussion about academic dishonesty. In addition, several demographic questions were added to the survey to gather information about the respondents.

The survey was administered to students in four different undergraduate programming courses: a first course in Java for information technology (IT) and information systems (IS) majors, a first course in Java programming for computer science (CS) majors, an introductory web page development course (HTML, CSS and JavaScript) for IT and IS majors as well as non-majors and a second course in Java for IS and IT majors. The survey was administered in class by a third party and was anonymous in that there was no identifying information on the survey. There were 155 respondents. All but five response was complete enough to use for analysis (n=150).

DATA ANALYSIS

Demographics of Respondents

The students that responded to the survey were mostly from three computing majors: Information Technology (46.7%), Information Systems (18.7%) and Computer Sciences (14.0%). The remainder of the respondents were from the College of Liberal Arts and Social Sciences or College of Science and Technology (19.3%), or undeclared or left their major blank (1.3%). Females accounted for 26.0% of the respondents, males 68.0% and the remainder did not identify their gender. Ninety-two percent (92%) of the respondents were age 24 or younger. Forty-seven point three percent (47.3%) of the respondents identified themselves as having a GPA of 3.0 or above.

Confirmatory Factor Analysis

To identify the underlying structure of the data and to reduce the number of variables in the analysis, a confirmatory factor analysis (CFA) was conducted on the 12 items on the survey corresponding to the programming, math, and English assignments for the survey items corresponding to perceptions on behaviors prior to education and after education confirming the category structure proposed in Table 1. Based on the initial fit of the CFA model, question 9 (making minor changes to an assignment submitted for a previous course and submitting it for the current course) was removed and question 6 (working together and submitting similar work) was moved from the category of unauthorized collaboration to copying part of an assignment. The rationale for removing question 9 was that it did not apply to the courses taught as they are primarily introductory in nature and also have vastly different assignments. The rationale for moving question 6 was a matter of fit. The statistics measuring fit for the CFA model improved when question 6 was moved from unauthorized collaboration and to copying part of an assignment. As the question can logically go in either category, the authors decided to include it in the copying part of an assignment category.

A summary of the final factors in the CFA are provided in Table 2. A summary of the statistics related to overall fit of the final model are provided in Table 3. The level of significance is greater than the recommended 0.05 for the programming assignments, chi-square divided by the degrees of freedom (chi-square/DF) is less than the recommended 3 in all cases, the normed fit index (NFI) is greater than the recommended 0.90 in all cases, the Tucker-Lewis index (TLI) is greater than the recommended 0.90 in all cases, the comparative fit index (CFI) is greater than the recommended 0.90 in all cases, the root mean
square error of approximation (RMSEA) is less than the recommended 0.80 indicating a close or reasonable fit for all but the math assignments and the standardized root mean squared residual is less than the recommended 0.10 for all but the math assignments (Kline 2005). Overall, based on the combined statistics, the model has an acceptable to good fit for all three types of assignments for the four factors identified in Table 3. Finally, reliability of each measure was assessed using Cronbach's alpha (Cronbach, 1951) and the values are provided in Table 4. In all cases, Cronbach's alpha is below the recommended level of 0.70 (Hair et al., 1998).

Table 2: Final four factor solution from CFA

<table>
<thead>
<tr>
<th>Category</th>
<th>Survey Questions Corresponding to CFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authorized Help</td>
<td>1 - 3</td>
</tr>
<tr>
<td>Unauthorized Discussion</td>
<td>4 - 5</td>
</tr>
<tr>
<td>Copying Part of an Assignment</td>
<td>6 - 8</td>
</tr>
<tr>
<td>Copying All of an Assignment</td>
<td>10 - 12</td>
</tr>
</tbody>
</table>

Hypothesis Testing

To determine whether or not students have different perceptions about what constitutes academically dishonest behavior for different types of assignments, a comparison of student perceptions for each category of behavior is made across the different types of assignments. More specifically, a comparison of programming assignment perceptions will be made to those of math and essay assignments. The results of the individual paired t-tests related to the hypothesis are presented in Table 5 and Table 6. Table 5 provides the results of tests conducted for the hypothesis comparing perceptions on the factors related to perceptions on programming assignments to those on essay assignments using a paired t-test (H1). Table 6 provides the results of tests conducted for the hypothesis comparing perceptions on the factors related to perceptions on programming assignments to those on math assignments using a paired t-test (H2).

Table 3: Fit statistics for final four factor solution from CFA

<table>
<thead>
<tr>
<th></th>
<th>Programming</th>
<th>Essay</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>47.603</td>
<td>68.846</td>
<td>96.373</td>
</tr>
<tr>
<td>Significance</td>
<td>0.137*</td>
<td>0.002</td>
<td>&lt; 0.000</td>
</tr>
<tr>
<td>Chi-square/DF</td>
<td>1.253*</td>
<td>1.812*</td>
<td>2.536*</td>
</tr>
<tr>
<td>NFI</td>
<td>0.959*</td>
<td>0.934*</td>
<td>0.925*</td>
</tr>
<tr>
<td>TLI</td>
<td>0.987*</td>
<td>0.955*</td>
<td>0.931*</td>
</tr>
<tr>
<td>CFI</td>
<td>0.991*</td>
<td>0.969*</td>
<td>0.952*</td>
</tr>
<tr>
<td>RMSEA</td>
<td>0.041*</td>
<td>0.074**</td>
<td>0.102***</td>
</tr>
<tr>
<td>SRMR</td>
<td>0.0481*</td>
<td>0.0618*</td>
<td>0.793***</td>
</tr>
<tr>
<td>Fit Assessment</td>
<td>Good</td>
<td>Acceptable</td>
<td>Acceptable</td>
</tr>
</tbody>
</table>

*a good fit, **a reasonable fit, ***a poor fit according to statistic
Table 4: Cronbach's alpha for final four factor solution from CFA

<table>
<thead>
<tr>
<th>Factor</th>
<th>Programming</th>
<th>Essay</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authorized Help</td>
<td>0.884</td>
<td>0.836</td>
<td>0.878</td>
</tr>
<tr>
<td>Unauthorized Discussion</td>
<td>0.826</td>
<td>0.747</td>
<td>0.875</td>
</tr>
<tr>
<td>Copy Part</td>
<td>0.859</td>
<td>0.787</td>
<td>0.871</td>
</tr>
<tr>
<td>Copy All</td>
<td>0.920</td>
<td>0.936</td>
<td>0.927</td>
</tr>
</tbody>
</table>

Table 5: Results of t-tests for H1 (programming assignments compared to essay assignments)

<table>
<thead>
<tr>
<th>Category</th>
<th>Questions</th>
<th>Difference in means</th>
<th>Standard Deviation</th>
<th>t-test statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authorized help</td>
<td>1, 2, 3</td>
<td>-0.40959</td>
<td>0.70865</td>
<td>-7.149</td>
<td>***0.000</td>
</tr>
<tr>
<td>Unauthorized discussion</td>
<td>4, 5</td>
<td>-0.26316</td>
<td>0.81774</td>
<td>-3.968</td>
<td>***0.000</td>
</tr>
<tr>
<td>Copy part</td>
<td>6, 7, 8</td>
<td>-0.50658</td>
<td>0.84916</td>
<td>-7.355</td>
<td>***0.000</td>
</tr>
<tr>
<td>Copy all</td>
<td>10, 11, 12</td>
<td>-0.05298</td>
<td>0.54173</td>
<td>-1.202</td>
<td>0.231</td>
</tr>
</tbody>
</table>

***significant at 1%, **significant at 5%, *significant at 10%

Table 6: Results of t-tests for H2 (programming assignments compared to math assignments)

<table>
<thead>
<tr>
<th>Category</th>
<th>Questions</th>
<th>Difference in means</th>
<th>Standard Deviation</th>
<th>t-test statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authorized help</td>
<td>1, 2, 3</td>
<td>0.02908</td>
<td>0.52122</td>
<td>0.681</td>
<td>0.497</td>
</tr>
<tr>
<td>Unauthorized discussion</td>
<td>4, 5</td>
<td>0.05369</td>
<td>0.83654</td>
<td>0.783</td>
<td>0.435</td>
</tr>
<tr>
<td>Copy part</td>
<td>6, 7, 8</td>
<td>0.13199</td>
<td>0.77578</td>
<td>2.077</td>
<td>**0.040</td>
</tr>
<tr>
<td>Copy all</td>
<td>10, 11, 12</td>
<td>0.01333</td>
<td>0.62137</td>
<td>0.263</td>
<td>0.793</td>
</tr>
</tbody>
</table>

***significant at 1%, **significant at 5%, *significant at 10%

The only significant difference in perceptions for programming versus math assignments is in copying part of the assignment (Table 6). It seems that students perceive copying part of a programming assignment as more unacceptable than one in math. There are differences in perceptions related to seeking authorized help, engaging in unauthorized discussion and copying part of an essay assignment as compared to a programming one (Table 5). Specifically, seeking authorized help for a programming assignment is more acceptable than for an essay; having unauthorized discussions is more acceptable for a programming assignment than for an essay; and copying part of a programming assignment is more acceptable than for an essay. The fact that students perceptions prior to taking the course are that copying part of a programming assignment and engaging in unauthorized discussions as more acceptable than for an essay assignment is problematic. The next logical question would be can we address this issue with education about what constitutes cheating as it specifically relates to programming and are there any changes in perceptions on graded essay and math assignments based on this education campaign.

DISCUSSION AND CONCLUSION

In an article in NetworkWorld (Marsan 2010) stated that 50% of the academic dishonesty cases at the University of Washington involved computer science students and 23% at Stanford. The article cites that reasons for cheating on programming assignments include (1) the availability of past solutions due to faculty not creating "new" assignments every semester as this is a difficult thing to do because of debugging requirements and (2) students thinking that there is only 1 correct solution so they do not realize solutions to the same problem have so much variation (so if a friend found the "right" answer and
there can only be one, how can I be caught?). This article highlights the fact that cheating in classes where programming is required is prevalent and needs to be addressed.

This study finds that students view programming assignments differently than other types of assignments which may be part of the reason cheating is so prevalent on programming assignments as found in the Marsan (2010) article. Students do have different perceptions regarding academic behavior for different types of assignments. They think that seeking authorized help is more acceptable on a graded programming assignment than on an essay. Of greater concern is that they view copying part of and participating in unauthorized discussions for a programming assignment as more acceptable behaviors than for an essay assignment.

To deal with the issue of unauthorized collaboration, Georgia Tech allows students to collaborate on programming assignments. When collaborating students have to sign a collaboration agreement and all collaborators on a project must be listed on the project. To ensure that each student learns the material, students must demonstrate their understanding of the program through an oral presentation of the code to a teaching assistant. In addition more of the grade come from the test and less from the homework (where students typically cheat more) (Marsan 2010). The attitude at Tech is that computing is best learned in a group and as long as students learn from another student's code, collaboration is OK.

Another way to address this issue of cheating in programming courses is for faculty to take the time to discuss issues related to academic dishonesty as it relates to programming as it makes a difference in perceptions. In a future paper, the authors plan to examine whether or not education can make a difference in perceptions.

REFERENCES


APPENDIX – SURVEY INSTRUMENT

You are working on a graded written essay for a class; your professor has told you this is an individual assignment. How acceptable are the following behaviors?

1. Asking the professor for help on the essay.
2. Asking a university provided tutor for help on the essay.
3. Reviewing similar essays in your textbook for ideas on how to write your essay.
4. Discussing ideas about the essay with a fellow student but writing the essays independently of each other.
5. Discussing ideas about the essay on an Internet news group, social networking site or blog.
6. Working together on the essay with a fellow student and submitting similar essays.
7. Copying a few sentences of another student's essay while adding a significant portion of your own work.
8. Copying a few sentences from the Internet or a written source while adding a significant portion of your own work.
9. Making minor changes to an essay you had previously written for another class and submitting it for this class.
10. Posting the assignment on an Internet news group, social networking site or blog asking someone to write the essay for you.
11. Hiring someone or asking a tutor to write the essay for you.
12. Copying another student's essay, making minor changes, and submitting it as your own work.

You are working on a graded programming assignment for a class; your instructor has told you that this is an individual assignment. How acceptable are the following behaviors?

1. Asking the professor for help on the program.
2. Asking a university provided tutor for help on the program.
3. Reviewing similar programs in your textbook for ideas on how to write your program.
4. Discussing ideas about the program with a fellow student but implementing the ideas independently.
5. Discussing ideas about the program on an Internet news group, social networking site or blog.
6. Working together on the program with a fellow student and submitting similar programs.
7. Copying a few lines of another student's program while adding a significant portion of your own work.
8. Copying a few lines of the program from the Internet or a textbook while adding a significant portion of your own work.
9. Making minor changes to a program you had previously written for another class and submitting it for this class.
10. Posting the assignment on an Internet news group, social networking site or blog and asking someone to write the program for you.
11. Hiring someone or asking a tutor to write the program for you.
12. Copying another student's program, making minor changes, and submitting it as your own.

You are working on a graded math assignment for a class; your professor has told you this is an individual assignment. How acceptable are the following behaviors?

1. Asking the professor for help on the assignment.
2. Asking a university provided tutor for help on the assignment.
3. Reviewing similar problems in your textbook for ideas on how to complete your assignment.
4. Discussing ideas about the assignment with a fellow student but implementing the ideas independently.
5. Discussing ideas about the assignment on an Internet news group, social networking site or blog.
6. Working together on the assignment with a fellow student and submitting similar work.
7. Copying a small part of another student's assignment while adding a significant portion of your own work.
8. Copying a small part of the assignment from the Internet or a textbook while adding a significant portion of your own work.
9. Making minor changes to an assignment you had previously completed for another class and submitting it for this class.
10. Posting the assignment on an Internet news group, social networking site or blog and asking someone to complete the work for you.
11. Hiring someone or asking a tutor to complete the assignment for you.
12. Copying another student's work, making minor changes, and submitting it as your own.