Getting students to earnestly do reading, studying, and homework in an introductory programming class

Alex Edgcomb[1,2], Frank Vahid[2,1], Roman Lysecky[3,1] and Susan Lysecky[1]
[1]zyBooks, Los Gatos, California
[2]Computer Science and Engineering, University of California, Riverside
[3]Electrical and Computer Engineering, University of Arizona
alex.edgcomb@zybooks.com, vahid@cs.ucr.edu, rlysecky@ece.arizona.edu, susan.lysecky@zybooks.com

ABSTRACT
Getting students to read and study before class, to be better prepared for lecture, or to enable a flipped classroom is a long-standing difficulty for teachers of introductory programming classes. Furthermore, getting students to do homework, consisting of small practice problems and questions, is also a long-standing difficulty without massive grading resources. And even then, preventing students from copying others' solutions is difficult as well. Today, the web enables new interactive learning material that is replacing past forms of textbooks and homework assignments, and students today commonly have access to needed devices and the internet. This paper provides data on student reading and homework completion rates for web-based interactive learning material we created that automatically records reading and homework activity by students. The data is for several thousand students at over 10 universities, for introductory programming classes in Java, Python, and C++. The data shows that, with an appropriate amount of awarded points, required-reading completion rate was 84%, and auto-graded homework completion rate was 75%, varying somewhat based on how many course grade points those items were worth. Students on average spent about 10 minutes reading each section, and about 3 minutes per homework problem, both appropriate amounts for those items. Furthermore, we developed measures of whether students were earnestly attempting the reading and homeworks, versus just "cheating the system" to get course grade points. We describe those earnestness measures in this paper. With proper design and amount of assigned work, 80%-90% of students earnestly did the reading and homework activities, even when no penalty existed for cheating the system, and fewer than 3% blatantly cheated the system to get their points.

CCS Concepts
• Social and professional topics~Code of ethics • Social and professional topics~CS1 • Social and professional topics~Student assessment • Applied computing~E-learning • *Human-centered computing~Empirical studies in interaction design*

Keywords
Student earnestness; Student cheating; Reading; Studying; Homework; Introductory programming; Learning content;

SIGCSE'17, March 8–11, 2017, Seattle, WA, USA.

ACM 978-1-4503-4698-6/17/03…$15.00.
DOI: http://dx.doi.org/10.1145/3017680.3017732

1. INTRODUCTION
Instructors of introductory programming courses commonly express the difficulty they have in getting students to prepare before lecture, such as reading a chapter or doing some basic programming activities. Such preparation is important when instructors wish to focus on harder concepts, do more examples, have students work on small problems together, and other modern teaching methods (including "flipped" classrooms). Students generally don't seem to read textbooks carefully (many don't even acquire textbooks today) and often don't watch pre-recorded videos carefully either. Requiring students to turn in homework before lecture or take quizzes at the beginning of lecture can help, but can consume scant grading resources and lecture time, and may create an unnecessarily-stressful environment for students. Thus, several years ago, we developed interactive web-based learning material for introductory programming, replacing a textbook as well as homework system. A key goal was to create "hands-on" material that provided an effective learning experience for students outside of lecture. The material has a "reading" component and a "homework" component. The reading material consists of some text, plus activities including (1) hundreds of animations of key concepts and (2) over a thousand interactive learning questions (short answer, multiple choice, true/false, matching), with hints and with explanations of right and wrong answers. Each topic is covered in a section, which includes a mix of such items. Student activity data, such as every submitted answer and click, is recorded in our database. An instructor can see the percentage activity completion per student, like watching an animation entirely or eventually correctly answering questions, and thus typically assigns some course grade points (e.g., 10% of the course grade) for completing "reading" by given due dates. Furthermore, the material includes "homework" problems, integrated throughout, consisting of short coding challenges (like "Read an integer from user input, and print that number doubled, then tripled"), which are auto-graded using several test cases, providing immediate feedback to the student.

Instructors often wonder if students will complete such activities. Thus, we analyzed completion data at over 10 universities, and report findings in this paper, showing high completion rates. Even more importantly, instructors often wonder if students will earnestly complete such activities, especially the reading activities which can be quickly completed just by clicking buttons. Thus, we developed various measures of earnestness that could be detected from the student activity data, and report findings in this paper, showing high earnestness. For example, we noted that a student could easily get all reading points of a section of material in about one minute, but instead that most students spend over 10 minutes per section.

The remainder of this paper describes our methods, and the completion and earnestness findings, for various reading and homework aspects of our material.

2. BACKGROUND AND RELATED WORKS

Use of high-quality learning material may be an important factor to increase student earnestness with completing assigned material. Newstead found that students are more likely to focus on learning from high-quality material than lower-quality, thus mitigating cheating [1]. Also, student motivation is higher when a student is studying to learn, rather than studying for points, and students are more inclined to cheat on classwork when course points are on the line [1][2][3]. Students are more inclined to cheat when the student was not sure what the question was asking [5][6][7][8]. These findings further emphasize the need for high-quality material to mitigate cheating.

Students and instructors sometimes have a different view on what constitutes cheating [7]. For example, many students reported as acceptable the checking of his/her own work by comparing to another student's work, whereas many instructors reported such behavior as cheating. Such differences in view may be due to a misunderstanding in the intentions by the student, but also suggests the need for better quality material that gives students the feedback they desire.

Automated homework systems are increasingly used in higher education to free up instructor time and provide immediate feedback to students [15]. Numerous companies have built such homework systems [4][9][10][11][12][13][14]. However, student earnestness on such automated homework systems seems yet to be analyzed. This paper analyses such automated homework systems.

This paper analyzes student activity from multiple zyBooks [4]. A zyBook is interactive learning material (intended to replace a textbook and homework system) that integrates many interactive activities, such as animations, tools, learning questions, and automated homework questions. Student activity is recorded, and instructors can view their student's activity in a dashboard.

This paper significantly extends our previous work [16], which analyzed student earnestness with short answer questions in a zyBook [16]. This paper extends by analyzing student earnestness of multiple choice questions, programming homework problems, and "What does this code output?" homework problems.

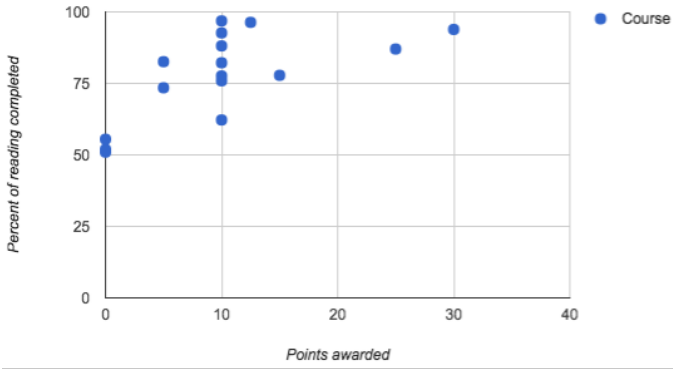
3. STUDENT ACTIVITY DATA

Anonymized student activity data was collected for analysis from zyBooks [4]. The courses in this analysis are introductory computer science courses at public and private institutions, including two-year and four-year institutions. The name of the courses and associated institutions are withheld to protect their privacy. Sections 4 and 5 include an explanation for how the specific courses were chosen for that section's analysis. As these were retrospective analyses, the students and instructors of the courses were blind to the purpose of this analysis.

4. READING FOR WEB MATERIAL

Reading web material includes reading some text and viewing figures, as in a traditional reading but typically less so, plus viewing animations, plus answering learning questions. Learning questions come in many types: true/false, multiple choice, short-answer, term matching, and more. In our material, learning questions also include explanations for all right answers, hints for wrong short-answer answers, and explanations for wrong multiple choice answers.

Figure 1: Students completion of the reading was relatively high, especially when points were awarded. Awarding more points had a slight positive increase on the percent of reading completed.



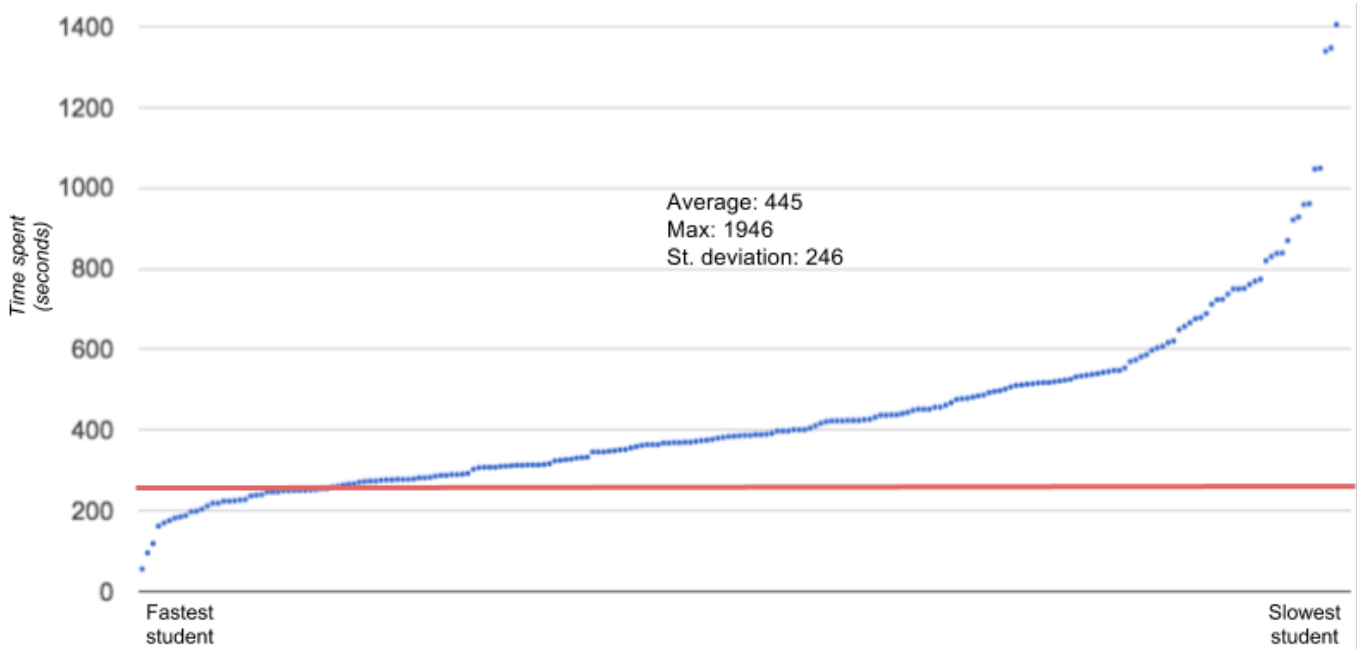
We measured the amount of reading completed versus the amount of points awarded for reading. Courses were identified by searching for course syllabi on the web. A course was included if the course syllabus indicated the number of points awarded for reading (including no points awarded), and if the course used a zyBook. This yielded 16 introductory programming courses with a total of 1,508 students. 3 courses awarded 0 points (127 students total), 2 courses awarded 5 points (257 students), 7 courses 10 points (710 students), 1 course 12.5 points (106 students), 1 course 15 points (57 students), 1 course 25 points (33 students), and 1 course 30 points (216 students). To account for different amounts of assigned work, the analysis considered the first five assigned chapters. Figure 1 shows that awarding more points tended to increase the percent of reading completed. Around 10 points awarded seems to be sufficient to convince students to read.

The above deals solely with completion. Instructors have asked: Do students earnestly complete the reading, or do they cheat the system to quickly earn points? We analyzed various earnestness measures: amount of time spent in the material, student behavior when answering short answer questions, and student behavior with multiple choice questions.

One measure of earnestness is spending an appropriate amount of time completing activities. An unearnest student would only spend enough time in a section to get full credit, whereas an earnest student would spend much more time. Note that the students were not aware that the earnestness of their activity would be analyzed. From the student's perspective, only completion of activities was important.

We compared the amount of time a student spends in a section to the fastest possible time. The fastest possible time was found by an undergraduate assistant completing the same section as fast as possible by just clicking buttons, without reading the subject matter.

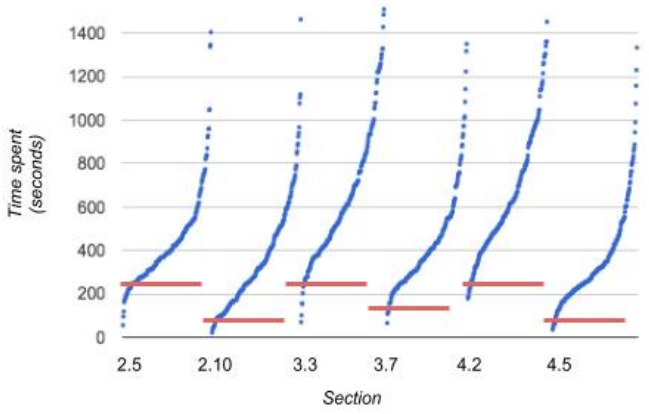
Figure 2: Amount of seconds that each student spent in section 2.5 of a C++ zyBook. Each blue dot is a different student. Red line is triple the fastest possible time. (130 students in total; one university's class)



We analyzed participation activity data to estimate the amount of time a student spends in a particular section. The analysis included 723 students across 3 introductory programming courses that used a zyBook: 202 C++ students, 149 Java students, and 372 Python students. The zyBook was required and instructors awarded points for completing the reading. The analysis included 5 to 6 sections per zyBook that had many participation activities. The analysis excluded chapter 1 to avoid dropped students and introductory material. The analysis also excluded students who did not complete the section.

Figure 2 and Figure 3 shows an example of student time spent in each analyzed section for one class. Each section is designed to require about 10-15 minutes from a learning student. Most students spent an appropriate amount of time in the section. The average time spent by a student was 587 seconds, whereas the average fastest possible time per section was 64 seconds. Thus, students spent about 10x more time than the bare minimum to get full credit.

Figure 3: Extends Figure 2 by showing 6 sections in a C++ zyBook. Most students spent an earnest amount of time in the section. Red line is triple the fastest possible time. (130 students in total; one university's class)

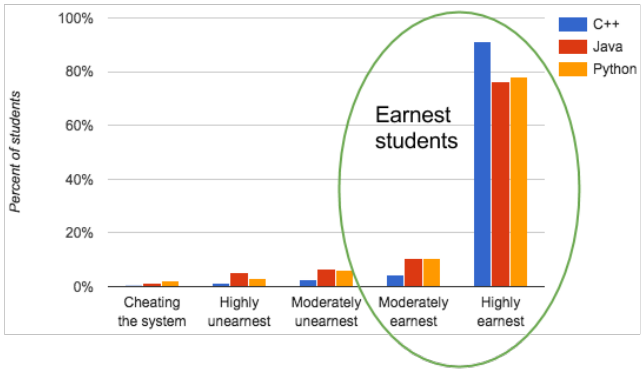


A second measure of earnestness is whether students attempted the short answer questions on the first try, as defined in previous work [16]. A short answer question includes a "Show the answer" button that students can click, intended for when students are stuck. An unearnest student might click the "Show the answer" button before attempting to answer the question, then simply type or copy-paste the answer to earn points. Submitting a blank answer or an all-whitespace answer was not considered an attempt.

We analyzed the same students and courses as for the above-described time spent analysis. This analysis included the first four assigned chapters of the course zyBook. The analysis included the earnestness categories: Highly earnest (>80% earnestness), moderately earnest (60-80% earnest), moderately unearnest (40-60% earnest), highly unearnest (20-40% earnest), and cheating-the-system (<20% earnest). As Figure 4 shows, about 90% of students were (highly or moderately) earnest; fewer than 2% blatantly cheated-the-system. These results are consistent with previous work [16].

A third measure of earnestness is the amount of time spent on multiple choice questions. A multiple choice questions includes the question, 2-3 choices, and an explanation per choice. An unearnest student might quickly finish the multiple choice question simply by rapidly clicking the choices until the correct choice is found. Whereas, an earnest student spends more time per multiple choice question to read the question and and choices. We determined 3 seconds to be a reasonable cutoff for unearnest behavior, found by looking at multiple choice activity data from 5 students who were highly earnest and 5 students who were cheating-the-system with short answer questions. The highly earnest students consistently spent more than 3 seconds per multiple choice question, whereas the unearnest students consistently spent under 3 seconds per multiple choice question.

Figure 4: About 90% of students earnestly worked through the short answer questions. Fewer than 2% blatantly cheated-the-system. (723 students in total; 3 courses/universities)



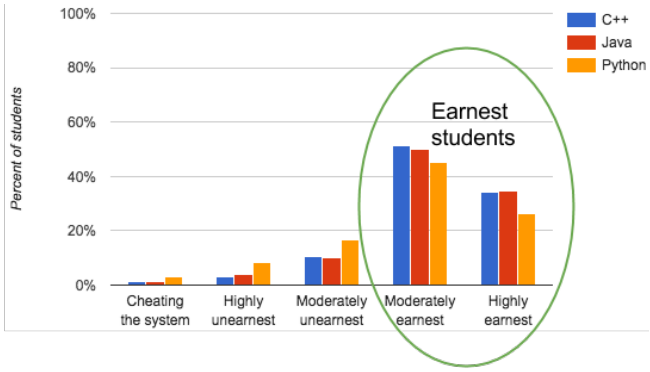
We analyzed the same students as the above-described time spent analysis, and included the first four chapters of the course zyBook. This analysis used the same earnestness categories as the short answer earnestness analysis. As Figure 5 shows, about 80% of students earnestly completed the multiple choice questions, and fewer than 3% cheated-the-system.

5. HOMEWORK

Homework differs from reading in that students cannot reveal the answer to a problem; instead students must correctly solve the problem to earn points. Homework for our web material includes auto-grading, immediate feedback, and no penalty for wrong answers. Examples include:

- Short "Write some code" challenges wherein the student is asked to complete code, typically by typing a few (typically 3-5) statements to achieve a particular goal. The code is run against test cases, which are immediately shown to the student.
- Short "What does this program output" challenges that typically include 5 difficulty levels per challenge, from easiest to hardest. Each level auto-generates a question. If the student gets a level right, then the student is given an explanation and moves to the next level. Else if wrong, the student is given the solution and the explanation, then given another auto-generated question of the same difficulty.

Figure 5: About 80% of students earnestly worked through the multiple choice questions. Fewer than 3% cheated-the-system. (723 students in total; 3 courses/universities)



We measured the completion rates of homeworks. Courses were identified in the same method as the above-described reading completion analysis. This method yielded 12 introductory programming courses with a total of 1,280 students. 2 courses awarded 0 points (98 students total), 1 course 5 points (193 students), 6 courses 10 points (606 students total), 1 course 12.5 points (106 students), 1 course 15 points (57 students), 1 course 25 points (33 students), and 1 course 30 points (216 students). To account for different amount of assigned work, the analysis considered the first five assigned chapters. Figure 6 shows that if points are awarded, then most students complete the homework. And, even without awarding points, many students still complete the homework.

One indicator of student earnestness with homework is the amount of time spent. An unearnest student will not spend time on the homework, instead copying solutions from a classmate or an online source (we easily found solutions posted online at a well-known site, for 36 of the 41 homework problems of a particular zyBook). We analyzed one such introductory C++ programming course with 262 students that used a zyBook. The analysis included 16 homeworks in the mid-chapters covered in the course. Time spent was computed by the time-between homework submissions, plus the time-between the first homework submission and that student's preceding activity submission. A

Figure 6: Most students completed the homework when points were awarded. Even without points, many students voluntarily completed the homework. Lower completion compared to participation activities is likely due to homework being more difficult. (11 courses/universities)

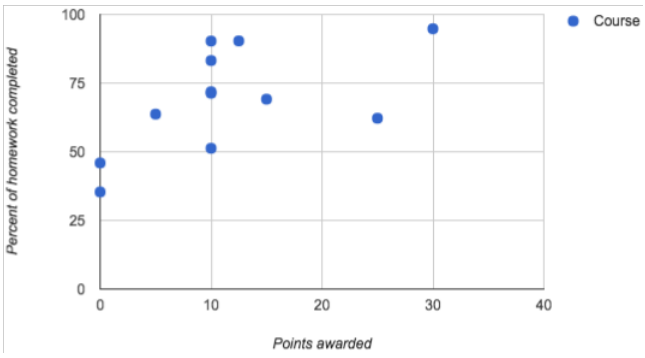


Figure 7: Most students spent an appropriate amount of time on the homework. (262 students across 16 homework problems; thus, 4192 homework problem submissions)

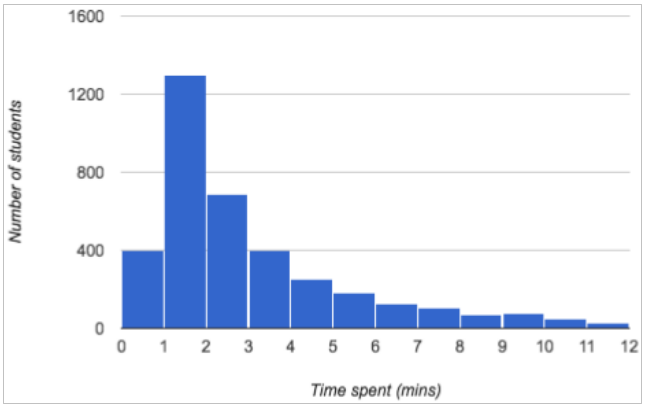
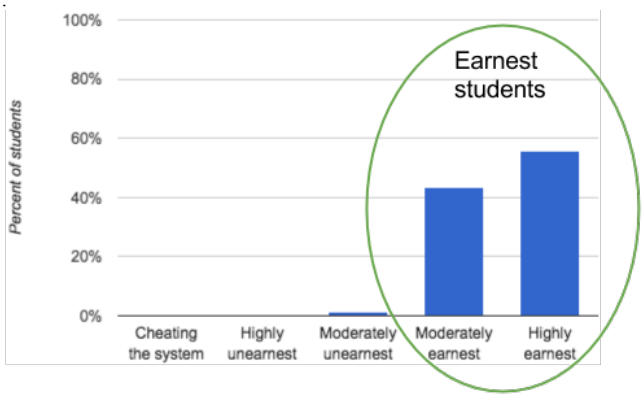


Figure 8: About 99% of students were earnest with the progression homework activity (259 students across 57 progression levels).



time-between greater than 10 minutes was excluded.

Each problem is designed to take about 1-3 minutes to complete for a student who understands the material. The average time a student spent on homework was 3.3 minutes. As Figure 7 shows, 90% of students spent at least 1 minute working on the homework. Note: Students completing in less than 1 minute are not necessarily unearnest; many problems could indeed be completed in under a minute by strong students. Thus, the earnestness % should likely be much higher here.

Another type of homework is a progression activity. An unearnest student might submit a wrong answer on purpose to see the solution, then use that knowledge to solve the next auto-generated question, whereas an earnest student will try to answer correctly first time. We analyzed 12 progression activities with a total of 57 levels, across 259 students in an introductory C++ course. The analysis counted the number of attempts until first time correct. So, if a student got the correct answer on the second try, then there were two attempts. This analysis used the same earnestness categories as the short answer earnestness analysis. As Figure 8 shows, 99% of students were earnest and 1% of students were moderately unearnest. Students averaged 1.4 attempts per level and spent on average 163 seconds per progression.

6. DISCUSSION

In general, the above analysis suggests that most students are not cheating the system to quickly earn points, even though no penalty exists for doing so. The most-clear evidence is that most students spend the expected amount of time doing the reading activities (about 10 minutes) and homeworks (about 3 minutes per problem). Additional evidence is that students are attempting short answer questions even though the answer can be shown with a single button click. Furthermore, time analysis shows that most students are legitimately trying on multiple choice questions rather than quickly clicking buttons.

The above does not imply that cheating does not exist. Clearly, cheating does exist and is quite common. The authors themselves, in the same classes using the C++ zyBook where students behaved earnestly, tend to catch 5-10% of the class cheating on programming assignments every term (sometimes even higher), resulting in stiff grade penalties. For some widely-used online learning materials from major publishers, we have indeed observed rampant cheating. Student comments include "Basically everyone I know cheated (popular tool name omitted)" or "We would work in groups and divide up the problems, thus increasing our allowed tries from 3 to 6, 9, 12, or more."

We believe earnestness is high on this particular material because students seem to realize the learning benefit and the efficient use of their time. End-of-term surveys at various universities support this belief both via numerical scores and student comments, like: “I wish I had this last semester”, “I think you’ve designed a fantastic resource”, “Thank you very much for your great textbook”, “Content is extremely well organized and the visuals are great tools”, “It truly helped me to understand on a deeper level”, “I really enjoyed it”.

From the authors' teaching experience, cheating tends to occur more on harder assignments (like weekly programming assignments), or when the simpler assignments are excessive or perceived as not helpful.

One limitation of this work is that we looked at retrospective data. Although numerous students were included, comparisons between populations were uncontrolled.

7. AUTHORING HIGH-EARNESTNESS CONTENT

Based on our experiences, we believe two key features improve student earnestness:

Efficiency:

- Useful activities: Any activity should be considered expensive, because student time is valuable. Is this item worth the cost? This mentality can permeate creation of *every* course activity. Given too many activities, students may give up and decide not to do any of them.
- Concise text: Students (and teachers) skim verbose text, whether in a book, syllabus, email, etc. Conciseness enhances communication and learning [17]. “If I had more time, I would have written a shorter letter” (often attributed to Blaise Pascal).

Understandability:

- Proper “tool” choice: No carpenter uses just a hammer. Teachers/authors should strive to use the right tool at the right time. Modern teaching tools include text, figures, animations, interactive questions, online tools (e.g., coding windows), videos, etc. Also, mixing tools aids student engagement.
- Small steps: Material often has gaps. Students need/appreciate if every item is carefully taught: “Baby steps”. Such steps may not impress professors, but material needs to be for students.

8. CONCLUSION

Across numerous universities and introductory programming courses, students earnestly completed both reading and homework of interactive web material. The keys to such earnest completion likely include awarding an appropriate amount of points, using properly designed content, and assigning a reasonable amount of work. Students spent about 10 minutes per section on average (as expected), completing 84% of required reading earnestly; fewer than 3% cheated-the-system. Also, students spent about 3 minutes per homework question (as expected), completing 75% of assigned homework earnestly.

Future work includes designing experiments to better target the key factors to best utilize interactive web material, and developing

techniques to make transparent student earnestness to both student and instructor without negatively impacting student learning.

9. ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation's Small Business Innovative Research (SBIR) program (1315094 and 1430537), and a Google faculty research award. We are grateful for their support. We are also grateful to Cyrus Hwang, Bailey Herms, and Kevin Nguyen for their contributions to the data analysis.

10. REFERENCES

[1] Newstead, Stephen E., Arlene Franklyn-Stokes, and Penny Armstead. "Individual differences in student cheating." Journal of Educational Psychology 88.2, 229, 1996.

[2] Jordan, Augustus E. "College student cheating: The role of motivation, perceived norms, attitudes, and knowledge of institutional policy." Ethics & Behavior 11.3, 233-247, 2001.

[3] Hutton, Patricia A. "Understanding student cheating and what educators can do about it." College Teaching 54.1, 171-176, 2006.

[4] zyBooks, www.zybooks.com, December 2016.

[5] Judy Shearda, Selby Markhama & Martin Dicka, "Investigating Differences in Cheating Behaviours of IT Undergraduate and Graduate Students: The maturity and motivation factors" Higher Education Research & Development Vol. 22, Issue 1, 2003.

[6] Avinash C Singhal, "Factors in Student Dishonesty" Psychological Reports 2011.

[7] Augustus E. Jordan, "College Student Cheating: The Role of Motivation, Perceived Norms, Attitudes, and Knowledge of Institutional Policy" Ethics & Behavior Vol. 11, Issue 3, 2001.

[8] Bernard E. Whitley, "Factors Associated with Cheating Among College Students" Research in Higher Education, Vol. 39, Issue 3, pp 235-274, 1993.

[9] Pearson, MasteringPhysics, http://www.pearsonmylabandmastering.com/northamerica/masteringphysics/, December 2016.

[10] Pearson, MathXL, http://www.pearsonmylabandmastering.com/northamerica/mathxl/, December 2016

[11] Wiley, WileyPlus, https://www.wileyplus.com/WileyCDA/, December 2016.

[12] TuringsCraft, CodeLab, http://turingscraft.com/, December 2016.

[13] WebAssign, WebAssign, http://webassign.net/, December 2016.

[14] McGraw-Hill, Connect, http://connect.mheducation.com/, December 2016.

[15] Scott Bonham, Robert Beichner and Duane Deardorff, "Online homework: Does it make a difference?", Phys. Teach. 39, 293, 2001.

[16] Yuen, J., Edgcomb, A., and Vahid, F. Will Students Earnestly Attempt Learning Questions if Answers are Viewable? Proceedings of ASEE Annual Conference, 2016.

[17] A. Edgcomb, F. Vahid, and R. Lysecky. Students Learn More with Less Text that Covers the Same Core Topics, Frontiers in Education Conference (FIE), IEEE, 2015.