# Survey of Carnegie Mellon Faculty on Student Attendance and Performance

**Introduction**

This study focuses on surveying members of the Carnegie Mellon faculty community in order to determine if there exists a relationship between attendance requirements and students' performance in classes. This topic is interesting because there is a large disparity in the way classes are structured across various departments of the university, and thus it is possible for two students of different majors to have entirely opposite classroom experiences. For example, many humanities courses are small and discussion based, while many science and math courses are large lectures composed of students from varying technical majors. In general, it would not be practical for instructors of such courses to require or take attendance because of the large class size. While focusing on determining if requiring attendance has an effect on students' performance, this survey will also, at the same time, note any other course details that may affect performance.

We hope that in completion of this survey, the Carnegie Mellon faculty and students will both benefit. By evaluating how course structure can affect performance, the University and faculty will gain insight on how to define structure and size limitations for their courses. The university should be interested in the information from this survey as they consider how many faculty members should be employed for different departments, as well as how they can create an environment where students can be most successful. In the same sense, students will also gain by better understanding the impact attendance can have on their grades and class performance.

Several of the sources we consulted found results indicating that class attendance and grades are correlated. Two studies that specifically looked at the relationship between attendance and scores on an exam found that students who attend class regularly have higher exam grades on average. Other related work has explored the effect of implementing mandatory attendance rules in classes, finding that such policies reduce absences and improve performance.

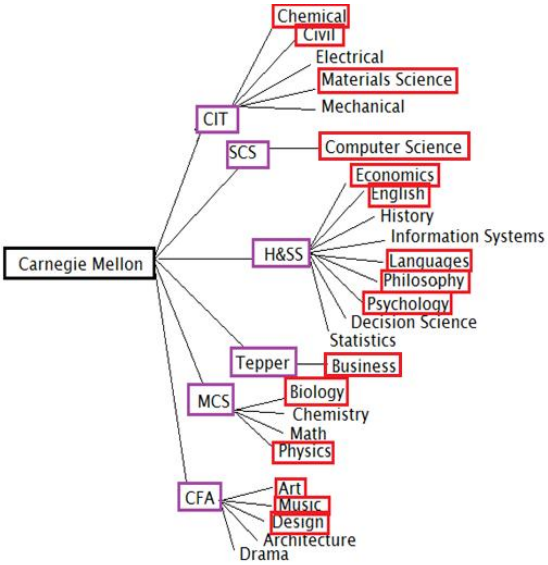
Our results indicated that there may be minimal effect of attendance policy on student grades at Carnegie Mellon. Along with this, we also find that attendance policy aside, classes where actual student attendance is high tend to have higher final grades. Finally, we see that smaller classes tend to have better student performance.

**Methods**

The target population of this study is Carnegie Mellon faculty members who teach undergraduates at the Pittsburgh campus. The population that we wish the make inferences about is Carnegie Mellon undergraduate students at the Pittsburgh campus.

In choosing the sample of the target population to survey, it was decided that a random sample would be taken in each of Carnegie Mellon's six major schools: College of Fine Arts, Carnegie Institute of Technology, College of Humanities & Social Sciences, Mellon College of Science, School of Computer Science, and Tepper School of Business. For each of these schools, we chose a random sample proportional to 50% of the number of departments (majors) within the school. We determined the listing of departments by visiting websites for each of six colleges, and randomly sampled from these listings. Making appropriate modifications for schools with an odd number of departments or schools with only one department, we assemble a sample of 15 departments. Therefore, our final survey design includes stratification by college and a clustered sampling of departments within each college.

| School | Number of Departments | Number of Departments Randomly Chosen (~50%) |
|---|---|---|
| CFA | 5 | 3 |
| CIT | 5 | 3 |
| H&SS | 9 | 5 |
| MCS | 4 | 2 |
| SCS | 1 | 1 |
| Tepper | 1 | 1 |




The graph above displays the breakdown our stratified sample. The departments highlighted in red indicate departments which were included in our sample.

*Sample Frame*

Originally, we had planned to use the Schedule of Classes website to find the courses within each department that had been chosen for our sample, and from those courses figure out which faculty taught an undergraduate Fall 2010 semester course. However, since we ran into problems obtaining the fall 2010 course list, we were forced to change our method for finding the e-mail addresses. We kept our initial sample design, but rather than using the sample frame of the directory after having found the faculty who taught undergraduate courses, we found the e-mail addresses of each department's faculty on the departmental website. To account for this tweak in our approach, we made sure to mention in the e-mail that only professors who taught an undergraduate 2010 fall semester course were to fill out the survey. If we had not specified this, we could have received responses from professors who taught only graduate courses, or only spring semester courses, which we were not interested in looking at in our survey. For the most part, the faculty did follow these directions aside from one or two faculty who may have misunderstood and filled out a response for the wrong semester.

*Survey*

We chose to create a new Gmail e-mail address for the purpose of sending out the e-mails, so that the responses wouldn't get lost in our individual e-mail accounts. This would also allow for easier monitoring of which e-mails we sent out, and which ones may have bounced back. After obtaining the e-mail addresses for each department, we sorted them into contact groups and sent out one e-mail to each departmental group. Within the body of our e-mail, we included a link to a Google Form survey we had

created. When each faculty filled out the survey, the response would then be added to our Google spreadsheet automatically.

The full listing of questions from the survey can be found in *Appendix 1*. The survey included questions regarding basic demographic course information (college, department, course number), class structure and policies (attendance policies and availability of lecture notes), as well as information about final grade distributions. To account for the construct of mandatory versus optional attendance, we included a definition for faculty to base their responses on. We defined attendance to be mandatory for any course where a student's grade will be subject to a direct measurable reduction if they do not attend class. This does not take into account the overall concept that students' grades will suffer if they miss class, but rather that some portion of their calculated grade is dependent on their attendance. This includes but is not limited to things such as attendance being taken and recorded in class, students only being allowed to miss a certain number of classes before facing a grade reduction, or having a lab or recitation where attendance is required in order to complete graded assignments.

*Sample Size and Margin of Error*

Because we ultimately decided to construct a sample of departments and sample all faculty members from those sampled department, we based our sample size calculation on the number of departments sampled as opposed to the number of faculty members. We initially considered it to be appropriate to sample approximately 50% of the departments in each stratum, and wished to determine if this would be appropriate. Because our response variable of interest is students' final grades, we based our calculations on estimates of student GPAs. Details of the following calculations can be found in *Appendix 3*. We first estimate the mean GPA across all colleges using the formula for stratified samples. We obtained estimates for the average GPA in each department by asking students of various majors and using appropriate weights for each stratum. We determined the estimate for the stratified sample mean to be 3.42.

Using the sample sizes for each stratum listed above, we calculate the sample variance for each stratum and the overall variance for the entire stratified sample. We determine the variance to be 0.005 and thus the standard deviation is 0.07. Because the School of Computer Science and Tepper School of Business only have one department, our sample variance in these strata was zero, which assisted in keeping our overall variance low. Using a 95% confidence interval, the margin of error is 0.1372. We considered these estimates for variance and margin of error to be quite acceptable, and thus proceeded with our intended sample of departments.

*Incentive*

Due to the fact that we didn't expect a high response rate because of how busy the faculty would be during this time, we provided an incentive in the form of a $40 gift card. This was not our original plan, however we were concerned that faculty would be less likely to respond around this time in the semester because We hoped that this would encourage faculty to fill out the survey because they would also have the option to write in their Andrew ID in order to be entered into a raffle, and thus would have a chance at a reward for filling out our survey. Despite this attempt to encourage more response, we still found our response rate to be quite low. In the future, we would recommend that any groups surveying the faculty do a test run to find out:

- Which e-mail subject headers are most effective and are more likely to be read
- What type of incentive is most appealing to faculty
- What e-mail name will be more likely to draw the faculty to click the e-mail, rather than ignore it

Had we had the time to test these factors before sending out our e-mails, we might have been able to observe a much higher response rate. We also feel that perhaps our non-response rate may have been low because many faculty did not follow the directions exactly for the survey – we asked that faculty fill out the survey once for each course that they taught in the fall semester, and since many faculty teach

multiple courses at once, we expected that more faculty would have had more than one response in our survey. We found that many faculty overlooked this request in the e-mail, or they may just not have wanted to take the time to fill out the survey several times. We also may have received a higher response rate if we had been able to afford a more valuable incentive, i.e. a $100 gift card, which may have been more appealing to the faculty.

*Follow-up*

We had initially intended to send out a preliminary e-mail asking faculty to participate in our survey, after which we would then follow-up in person for the faculty who hadn't responded. After e-mailing over 500 faculty members, we realized that this plan needed to be changed. Considering the amount of non-response that we did end up having, it was infeasible for us to visit several hundred faculty members in person. We changed this aspect of our design, and instead decided to conduct our follow-ups by e-mail. We sent out a reminder e-mail after about 4 days to ensure that those faculty members who hadn't responded would remember about our survey, in the case that our e-mail got overlooked in their inbox. Our follow-up e-mails resulted in an increase in response rate, although our overall non response still remained quite low.

**Problems and Solutions**

Throughout the process of completing this study there were several unforeseen problems that arose which forced a change in the original design. The first of these was a difficulty in attaining the list of faculty who taught courses in the Fall 2010 semester. Originally the plan was to use the list of courses and faculty to send only those who taught undergraduates the survey via email. Unfortunately, after 2 weeks of searching for the information we were informed that the list is not general public information, which we initially thought it would be. Since we were unable to attain the information in a timely matter we had to change our design slightly so as to utilize the information we could get publicly. Our solution was to use the available emails on each departmental website and send the survey to the entire listed faculty. While this ultimately yielded the results we wanted, it was not ideal as it forced us to email faculty that either did not teach a course in Fall 2010 or did not teach an undergraduate course.

nice discussion

The change in design resulted in a problem with non-response. Since the number of faculty emailed increased due to the design adjustments our response rate appears very low (approx. 11%). Unfortunately, without a list of faculty who fit the original criteria, we have no way of knowing exactly how many eligible units we had. The cause for the low response rate could have been caused by the fact that many of the faculty emailed did not fit the criteria and therefore would never have been contacted with initial design of the survey.

In our attempt to fix the non-response we encountered another, unanticipated, problem: spamming. Our original plan to deal with non-response was to send out a follow-up email within 24 – 48 hours after the initial email was sent out. Unfortunately, since we used a general Gmail account to send out the email, and the number of faculty we emailed the survey to increased (519 in total), Gmail tagged our account as spammers and refused to allow us to send out our follow-up email until several days after our first one was sent out. This meant that we were unable to deal with non-response as quickly as we would have liked, and could have further contributed to our low response rate.

**Results**

The following table summarizes the responses we obtained from each college and department.

| College | Department | Number of Responses | Total Responses |
|---|---|---|---|
| CIT | Chemical Engineering | 3 | 12 |
| | Civil and Environmental Engineering | 6 | |
| | Materials Science Engineering | 3 | |
| CFA | Art | 1 | 3 |
| | Music | 1 | |
| | Design | 1 | |
| H&SS | Economics | 0 | 22 |
| | English | 5 | |
| | Modern Languages | 8 | |
| | Philosophy | 4 | |
| | Psychology | 4 | |
| | Statistics | 1 | |
| Tepper | Business | 6 | 6 |
| MCS | Biology | 7 | 12 |
| | Chemistry | 1 | |
| | Physics | 4 | |
| SCS | Computer Science | 4 | 4 |

Our responses included 37 courses with mandatory attendance, and 22 courses with optional attendance. The following table shows a summary of the percentage of students' grades which is dependent on attendance. This is defined to be the amount attendance is weighted in the calculation of final grade in the course.

| Percentage of Final Grade Dependent on Attendance | Number of Classes |
|---|---|
| 0 – 10% | 18 |
| 10 – 20% | 7 |
| 20 – 30% | 7 |
| 30 – 40% | 4 |
| Other | 26 |

Of all the classes surveyed, 47 of them had lectures, 29 had recitations, and 14 had labs. Instructors could select more than one type of class meeting type in their responses.

23 of the classes were lectures by a professor, 13 were discussion based, and 23 were a combination of both lecture and discussion.

When asked about whether or not notes were available, 34 responses indicated that lecture notes were made available to students, and 25 responded that they were not available. For those where notes were available, for 10 of the classes notes were available in full text form, and 24 classes had notes where attendance was necessary for students to have them complete.

In examining the responses to our survey, we decided to focus on determining if there were relationships between students' final grades and the following factors:
-Whether or not the course has mandatory attendance
-Percentage of actual student attendance in class

would be interesting to look at relationship of this with mandatory attendance...

-Size of class

**Hypothesis***: Attendance mandatory courses will have higher mean final grades.*
Our main goal was determining whether or not there is a relationship between students' final grades and the class attendance policy. In Figure 1 we can see that there are not any obvious differences between the shape, center, and spread of attendance mandatory classes and optional classes. The median is about the same for both groups, although the spread of the distribution for attendance mandatory courses is a bit larger. We would ultimately conclude that there is not strong evidence for a relationship between grades and attendance policy. We had expected that there would be a larger difference between grades for the two groups.

what would a 2-sample t-test tell you?

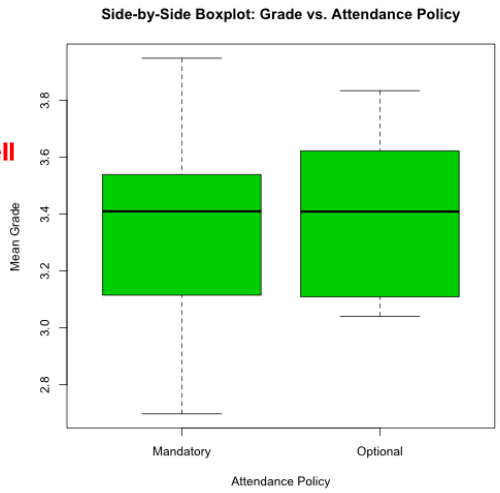



Figure 1. Boxplot of Mean Final Grades by Attendance Policy

The table below shows summary statistics for the distribution of final grades based on whether or not the class has mandatory attendance.

| **Mean Final Grade** | Minimum | Median | Mean | Maximum |
|---|---|---|---|---|
| All Classes | 2.697 | 3.409 | 3.363 | 3.947 |
| Attendance Mandatory | 2.697 | 3.409 | 3.356 | 3.947 |
| Attendance Optional | 3.040 | 3.408 | 3.394 | 3.833 |

**Hypothesis:** *Classes with a higher percentage of students attending class will have higher final grades*
Another relationship that we wished to investigate was between students' final grades and the actual percentage of students who attended class. In Figure 2 we can see 3 boxplots that show the distributions of grades based on 3 different attendance percentage categories. Based on the graph, we can see that those classes in which 81-100% of the students attended class had a much higher median final grade than the 2 lower categories. The max range is also greater than the other 2 categories. This would suggest that there is a positive relationship between a students' final grade and the actual percentage of students who attend class.

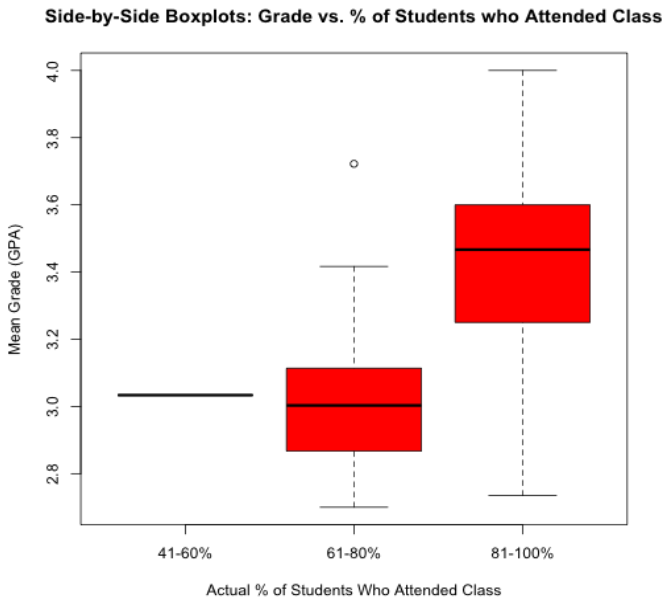



Figure 2. Boxplots of Final Grades by Class Attendance

The table below shows summary statistics for the mean final grades based on the percentage of students who normally attend class. Note sample size for the 41-60% group was only one course.

| **Mean Final Grade** | Minimum | Median | Mean | Maximum |
|---|---|---|---|---|
| 41-60% Attendance | 3.034 | 3.034 | 3.034 | 3.034 |
| 61-80% Attendance | 2.700 | 3.003 | 3.057 | 3.722 |
| 81-100% Attendance | 2.735 | 3.467 | 3.433 | 4.000 |

**Hypothesis:** *Mean final grade decreases as class size increases.*
We thought that perhaps we would see slightly lower mean final grades for larger classes for several reasons: first, the fact that a larger class would have a wider variety of students, thus there may be a greater amount of students who do poorly. Second, since the smaller classes tend to be more attendance/discussion based and place less emphasis on exams, there may be more A's in these type of classes simply because it is more difficult to fail a discussion. There is also the idea that smaller classes would have more attendance because it is more obvious when a student does not show up in a small class, hence they would try to increase their attendance and participation, and therefore their grade if the grade is heavily based on attendance/participation. All of these could be factors that play into the relationship between class size and mean final grade.

In Figure 3 we examine this relationship graphically. We construct a scatterplot, including a lowess smoother to better illustrate the relationship between the data. The data seems to be pretty evenly spread out from a mean final grade of 2.8 to 4.0 within class sizes that are 100 or below. The lowess smoother shows somewhat of a negative relationship, with regards to mean final grade decreasing the larger the class size. Since we did not get as many responses from professors who taught larger classes above 100 students, it is difficult to confirm whether this actually holds true for most of the classes at Carnegie Mellon. However, from the data points that we did receive, we do see suggestions based on the data points decreasing that there may be a relationship between class size and mean final grade.

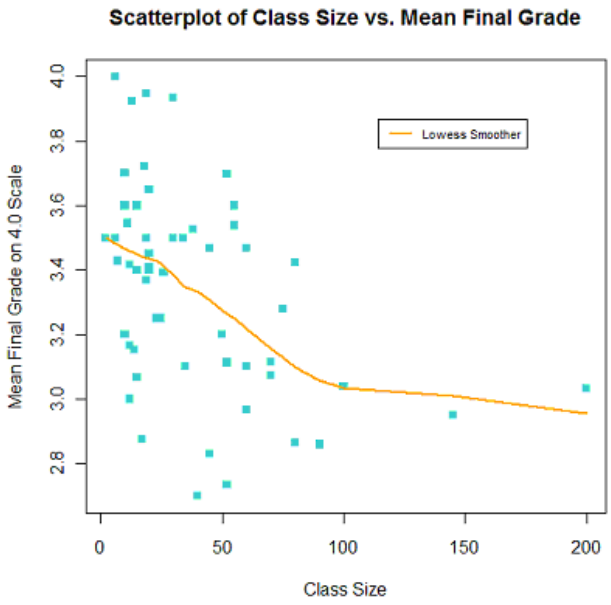



Figure 3. Scatterplot of Mean Final Grade versus Class Size

The table below gives a summary of the minimum and maximum for each variable (class size and mean final grade), as well as the median and mean.

| | Minimum | Median | Mean | Maximum |
|---|---|---|---|---|
| Class Size | 2.00 | 25.5 | 39.24 | 200.00 |
| Grade | 2.70 | 3.40 | 3.334 | 4.00 |

These are the minimum, maximum, and mean across all colleges and departments that we sampled. We see a mean of 3.34 for the mean final grades among the courses, and a mean of 39.24 for class size, which seems somewhat small. This may be because we received more responses from professors conducting smaller courses than larger courses that are over 100 students. We see that the range for class size ranges from 2 students to 200, while the mean final grade ranges from 2.7 to 4.0. There are many other factors that differ between the classes that could also influence the mean final grade, however we were interested in seeing whether there was any direct relationship between class size and mean final grade.

**Conclusion**

By analyzing the course demographics and grade distributions provided by the survey participants, we are able to draw a number of conclusions regarding factors that affect course performance at CMU.

In regards to our first hypothesis on attendance policy and course performance, we are able to conclude that our data showed results opposite to what we predicted. Before any data was obtained, it was thought that courses with mandatory attendance would have higher grade averages, due to the fact that students would be required to attend class. After analyzing the results, we concluded that attendance policy actually had no effect on course performance, because with both policies, the average grade for the class was about the same.

interesting contrast

In regards to our second hypothesis on approximate class attendance and course performance, the results obtained show an implication parallel to our predictions. Prior to conducting the study, we hypothesized that greater class attendance would result in better class performance, going off the idea that attending

class should help students do better overall. After analyzing results, we conclude that while attendance policy may not be influential on class performance (hypothesis 1), actual attendance levels definitely are.

Finally, in regards to our last hypothesis, the findings showed that as class size increases, mean class grades do in fact decrease. In agreement with our predictions, we concluded that class size has a correlation with class performance, despite the fact that there are many side factors that could affect this outcome.

nice, unanticipated finding...

By taking into consideration the findings from our survey, it seems that not only the university, but also faculty and students, will be able to benefit. If the university can take into account the fact that class size has a negative impact on class performance, they may realize that having more faculty to teach courses with smaller class sizes will be helpful for students' learning and performance. Along with this, faculty themselves can learn that attendance policies do not actually heavily affect grades. Enforcing mandatory or optional attendance for their courses does not actually mean more students attend class and for this reason, it does not seem to have an impact. Further analysis on the correlation between attendance policy and actual course attendance may be beneficial in considering this point. Finally, our findings regarding actual class attendance will benefit students in that by attending class, average class grades to increase. This will let students understand that while it may seem inconvenient or pointless to attend class, there is evidence that greater class attendance results in higher class performance. On the whole, we hope that this survey on student attendance and class performance will provide useful and beneficial information to those mentioned above in terms of structuring and designing future courses.

*Appendix 1: Questionnaire*

1. College:
2. Department/Major:
3. Course Number:
4. Size of Class:
5. Attendance Mandatory or Attendance Optional?
6. Is this class lecture only? Are their recitations or labs?
7. What percentage of students enrolled attend class regularly?
   **0-20% 21-40% 41-60% 61-80% 81-100%**
8. Is this a core or major required course?
9. Class Structure: Discussion based or lecture based?
10. Are notes or lectures available for students to view online?
Are they complete or partial notes? (Do students need to be in class to fill in certain sections?)
11. If attendance mandatory:
What percentage of student's grade is dependent on attendance?
   **0-10% 10-20% 20-30% 30-40%**
12. Grade distribution for a previous semester:
   Number of students receiving an A:
   Number of students receiving a B:
   Number of students receiving a C:
   Number of students receiving D/F:
   Mean Final Grade(if applicable):
Is this distribution consistent with previous semesters?

*Appendix 2: Consent Form*

**Form for Online Consent**

This survey is part of a research study conducted by Christopher Chang, Kelly L. Chang, Emily Boncek, and Stephanie Sindler at Carnegie Mellon University.

The purpose of the research is to determine if there is a relationship between whether or not a class has mandatory attendance and students' performance in the class. This topic is interesting because there is a large disparity in the way classes are structured across various departments of the university, and thus it is possible for two students of different majors to have entirely opposite classroom experiences. More specifically, many humanities courses are small and discussion based, while many science and math courses are large lectures composed of students from varying technical majors. In general, it is not practical for instructors of such courses to require or take attendance because of the large number of students. This survey is interested in determining if requiring attendance has an effect on or can improve students' performance in classes.

**Procedures**
Participants will be asked to fill out a survey through e-mail or a face-to-face survey per the participant's request, describing information about courses that they've taught the past semester.

**Participant Requirements**
Participation in this study is limited to individuals age 18 and older. Participants must be a full-time faculty member teaching at Carnegie Mellon University.

**Risks**
The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life or during other online activities.

**Benefits**
There may be no personal benefit from your participation in the study but the knowledge received may be of value to humanity, particularly to the Carnegie Mellon campus community.

**Compensation & Costs**
There is no compensation for participation in this study.

There will be no cost to you if you participate in this study.

**Confidentiality**
The data captured for the research does not include any personally identifiable information about you. Your IP address will not be captured.

By participating in this research, you understand and agree that Carnegie Mellon may be required to disclose your consent form, data and other personally identifiable information as required by law, regulation, subpoena or court order. Otherwise, your confidentiality will be maintained in the following manner:

Your data and consent form will be kept separate. Your consent form will be stored in a locked location on Carnegie Mellon property and will not be disclosed to third parties. By participating, you understand and agree that the data and information gathered during this study may be used by Carnegie Mellon and published and/or disclosed by Carnegie Mellon to others outside of Carnegie Mellon. However, your name,

address, contact information and other direct personal identifiers in your consent form will not be mentioned in any such publication or dissemination of the research data and/or results by Carnegie Mellon.

**Right to Ask Questions & Contact Information**
If you have any questions about this study, you should feel free to ask them by contacting the Principal Investigators now at klchang@andrew.cmu.edu, eboncek@gmail.com, cjradinfo@gmail.com, or ssindler@andrew.cmu.edu. If you have questions later, desire additional information, or wish to withdraw your participation please contact the Principle Investigators by e-mail in accordance with the contact information listed above.

If you have questions pertaining to your rights as a research participant; or to report objections to this study, you should contact the Research Regulatory Compliance Office at Carnegie Mellon University. Email: irb-review@andrew.cmu.edu . Phone: 412-268-1901 or 412-268-5460.

The Carnegie Mellon University Institutional Review Board (IRB) has approved the use of human participants for this study.

**Voluntary Participation**
Your participation in this research is voluntary. You may discontinue participation at any time during the research activity.

I am age 18 or older. Yes No
I have read and understand the information above. Yes No
I want to participate in this research and continue with the survey. Yes No

*Appendix 3: Sample Size Calculation*

In performing our calculation of sample size, we first estimated the overall mean GPA across all of the departments in our sample. We used the equation for stratified samples $y_{str} = \sum^{H} W_h y_h$, where $W_h$ is the weight for each stratum, or the number of departments in the stratum divided by the total number of departments in the population. $y_h$ is the mean GPA for each stratum. We obtained estimates of the mean GPA in each department by surveying individuals whose majors were in those departments.

We initially determined that we wanted to sample approximately 50% of the departments in each college. Our sampling percentage ended up being higher than 50% for some colleges such as those with odd numbers of departments in which case we rounded up, or for the college which had only one department.

To calculate our variance and margin of error, we calculated the sample variance in each stratum using the formula $s^2_h = 1/(n_h - 1) \sum^{nh}(y_{hi} - y_h)^2$ where $y_{hi}$ is the mean GPA for each department and $y_h$ is the mean GPA for the associated college.

To calculate the overall variance for our stratified sample $Var(y_{str}) = \sum^{H} W^2_h(1 - n_h/N_h)*(s^2_h/n_h)$, where $n_h$ is the number of sampled departments in each stratum, and $N_h$ is the total number of departments in the stratum.

We estimated our 95% margin of error by multiplying our standard deviation, obtained from the square root of the variance, by 1.96.

*Appendix 4: Survey Invitation E-mail*

**NOTICE: If you did not teach an Undergraduate course in the past Fall 2010 semester please ignore this message. Thank you!**

Dear Carnegie Mellon Faculty,

You have been selected in a 1-stage clustered random sample to take part in an academic, student survey.

This survey is being conducted for the Carnegie Mellon undergraduate class 36-303: Sampling, Survey, and Society.

We would greatly appreciate it if you could take a few minutes to fill out a short survey regarding the courses you taught in this past Fall 2010 semester. Our hope is to use this information to determine whether or not there is a relationship between course structure and student performance.

Provided below is the link to our survey. It should only take approximately 5 minutes. Our deadline for data collection is April 22nd. **Upon completion of the survey you will be asked for your AndrewID in order to be entered in a raffle for a $40 Union Grill Gift Card.** If you'd rather remain anonymous then you may choose to not participate in the raffle.

If you taught multiple undergraduate courses in the 2010 fall semester, please fill out the survey for each applicable course:

https://spreadsheets.google.com/embeddedform?formkey=dGhZMXM2RlBoVXpwZWRRbXN6X3F1NUE6MQ

Your answers are confidential, and will only be used for the purpose of our group project for 36-303 and the results will be presented at Carnegie Mellon's Meeting of the Minds. Thank you very much for your participation, and please let us know if you have any questions or concerns at 303teamc@gmail.com

Sincerely,
Kelly Chang
Emily Boncek
Christopher Chang
Stephanie Sindler