Work in Progress: Analysis of the impact of office hours on graded course assessments

Natalia Ozymko, University of Illinois at Urbana - Champaign

Natalia Ozymko is a rising senior majoring in Computer Science with a minor in Spanish at the University of Illinois at Urbana-Champaign (UIUC). She is interested in helping students master advanced topics in Computer Science and building new technologies to improve people's lives. She was awarded the Scott Fisher Outstanding Course Assistant award, and has worked under the direction of multiple faculty members assisting in teaching both Data Structures and Systems Programming.

Matthew Allan McCarthy, University of Illinois at Urbana - Champaign

Matthew McCarthy is a Junior in Mathematics at The University of Illinois at Urbana-Champaign. He likes to do programming projects and data analysis in his free time. For this past year he has been working under different faculty members in both research and software development. He hopes to make the world amazing with his work.

Prof. Wade Fagen-Ulmschneider, University of Illinois at Urbana - Champaign

Wade Fagen-Ulmschneider is a Teaching Associate Professor of Computer Science at the University of Illinois at Urbana-Champaign (UIUC). With a passion for data, he teaches thousands of students each year in his courses on Data Structures, Data Visualization, and Data Science. He was selected as one of the National Academy of Engineering's Frontiers of Engineering Education scholars, awarded the Collins Award for Innovation Teaching, and has been consistently ranked as an excellent instructor by his students for the past ten years. His work on data visualizations has been used by governors of multiple states, featured by websites including Popular Mechanics and The Verge, and has been viewed by millions of readers.

Prof. Karin Jensen, University of Illinois at Urbana - Champaign

Karin Jensen, Ph.D. is a Teaching Assistant Professor in bioengineering at the University of Illinois at Urbana-Champaign. Her research interests include student mental health and wellness, engineering student career pathways, and engagement of engineering faculty in engineering education research. She was awarded a CAREER award from the National Science Foundation for her research on undergraduate mental health in engineering programs. Before joining UIUC she completed a post-doctoral fellowship at Sanofi Oncology in Cambridge, MA. She earned a bachelor's degree in biological engineering from Cornell University and a Ph.D. in biomedical engineering from the University of Virginia.

Karle Flanagan, University of Illinois at Urbana - Champaign

Karle Flanagan is a Senior Instructor of Statistics at the University of Illinois at Urbana-Champaign. She has taught introductory statistics to thousands of students at UIUC since Spring of 2014. She also serves as the MS advisor for the statistics department. In 2018, she was awarded the Illinois Student Government's Teaching Excellence Award and in February of 2019, she also won the highest level teaching award at UIUC, the Campus Award for Excellence in Undergraduate Teaching. She completed her undergraduate degree in mathematics, with a minor in secondary education. She previously has taught mathematics and worked as a statistician in the insurance industry. Along with teaching, she is currently working on course development for other advanced statistics courses and data science courses using Python. Her research interests include online education, optimizing efficiency in office hours for large classes, and active learning methods for undergraduate statistics instruction.

Work in Progress: Analysis of the impact of office hours on graded course assessments

0. Abstract:

In this work, we used log files gathered from an online queueing system and combined those logs with the scores students earned on graded assessments. With data from four sections across two semesters of a large sophomore-level computer science course, this work is the largest known observational analysis of the impact of office hour attendance on graded assessments. This work in progress begins this analysis by exploring the relationship between office hour attendance and graded assessments over a full academic year in a large Data Structures course (n=1,238 students).

Our initial findings suggest that there are several relationships that warrant further exploration. The first major finding is that office hour attendance provides a significant increase to a student's score on upcoming graded homework; however, it does not provide a significant boost to a student's score on upcoming exams. The second major finding is that the overall impact on a student's course grade by attending office hours decreases the closer that student attended office hours relative to an assignment due date or exam date.

Our work outlines the statistical techniques used in our analysis, explores differences between various sections of a course across two semesters, and provides an outline of recommended changes for how office hours are run based on lessons learned from this analysis. In the future, we hope that this will lead to improved learning, which will improve students' mastery of the material and problem-solving abilities.

1. Introduction:

Office hours are a common feature of many university courses. A traditional model for hosting office hours involves a student showing up to a specific location, asking the instructor one or more questions, and then leaving office hours. It is common for some larger courses to hold "group office hours", where the instructor would generally answer questions in a group setting. While both of these types of office hours have worked for different types of courses, they can be difficult in large courses because of a dramatic increase in the volume of questions asked. Furthermore, the diversity of student solutions to programming assignments can make group office hours challenging. To solve this, we have adopted an open source Queue. [1] The Queue allows students to add themselves to the office hours "queue" as they show up, and provides a clear student ordering for course staff.

For the past two years, nearly a dozen University of Illinois courses have adopted the use of a web- and phone-based queueing system for office hours. Through the use of this queueing system, over 8,000 students attended office hours at least once, and over 70,000 individual

questions were facilitated. The use of this system was initially designed to streamline the process of students writing their names on a whiteboard to get help. Although this worked for smaller courses, larger courses require more structure for office hours. The Queue allows course staff to more effectively help students while collecting data about its users and the question they ask. With its continued use, the data collected provides powerful course analytics that could be used to improve learning and the student experience.

One benefit to the Queue system, compared to traditional office hours, is an increase in data collected. Analyzing this data can provide insights into courses, such as what times might need more course staff scheduled or what assignments students are struggling with. This paper analyzes a complete academic year (AY 2018-2019) of data from a sophomore-level Data Structures course that utilizes the Queue for all office hours held by the course.

The core contribution of this paper is an initial analysis of graded assessments -- through the analysis of anonymized student grade data for each course assessment and the impact the attendance of office hours had on graded assessments. The assessments included programming assignments (7 total), lab assignments (14 total), quiz and exam grades (8 total), and weighted averages of those assessments (e.g., final course grade).

Through the analysis of data on utilizing office hours and course grades, this study found: **(1):** Students who attended office hours typically earned higher scores on graded assignments, but not on proctored exams, **(2):** Students who sought help on assignments closer to the due date earned slightly lower final grades than those who got help earlier, and **(3):** Student use of staffed office hours follows the Pareto 80-20 rule where 80% of the staff time is spent answering questions from 20% of the students. [1, 2]

This paper provides background about the queue system and data sources used (Section 2), reviews related work about this topic (Section 3), analyzes the results (Section 4), and finally, discusses future work using the data in this paper combined with other sources of data (Section 5).

2. Background and Methods: Queue Software, Queue Usage Logs, and Gradebook Data

The Queue is a web-based application that facilitates student and instructor interactions during office hours by allowing students to add themselves to a virtual line when they have a question to ask the course staff. Course staff help students on the Queue according to a first-come first-served system, where the student at the top of the Queue is always selected to be helped next. This system makes helping students during office hours organized and efficient, as there is now a clear ordering for course staff, and students can estimate how long they will wait until their turn.

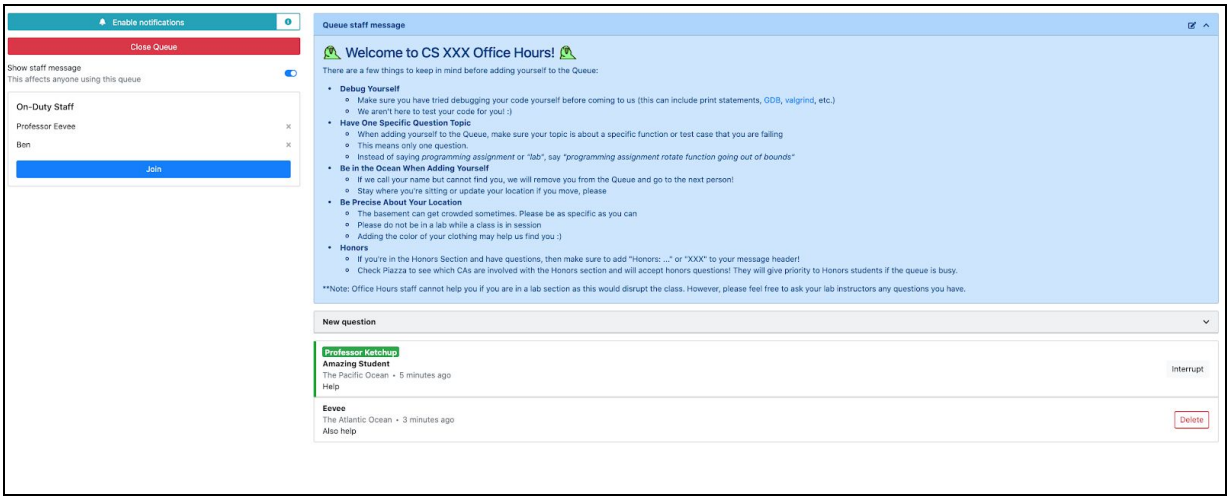

Figure 1. A screenshot of the web-based Queue interface on a laptop computer. The Queue interface shows students the current on-duty course staff helping to answer questions (left area), a course-specific message that is provided to everyone asking questions (large shaded box), and a list of students currently on the Queue (list shown below the shaded box).

The Queue software logs each event that occurs on the queue. To evaluate the impact of office hours, this work extracted the timestamps from Queue server logs of when a student first asked a question on the Queue, when each student question began being answered by course staff, when each question was marked as having been answered, if the student left the Queue before their question was answered, and other data. All student and staff identifiers were anonymized.

This work specifically analyzed the office hours data of a sophomore-level Data Structures course at The University of Illinois called CS 225: Data Structures. Throughout the academic year-long analysis period (Fall 2018 - Spring 2019), the Queue recorded that 759 of the 1,238 enrolled students (61%) made use of office hours at least once during the academic year.

To study the impact of office hours on course grades, the grades of students on all assessments in CS 225 was added to the anonymized queue logs for this analysis. CS 225 grades are split up between programming assignments, labs, exams, and a final exam. Programming assignments are two-week long programming assignments that are released on Tuesday every other week. All exams and the final exam are proctored exams held in a computer-based testing center.

The Queue facilitates office hours that are run by instructors, teaching assistants, and undergraduate course aides, collectively referred to as “course staff”. Students work anywhere within a specified area (e.g. a large atrium area of an academic building) and add themselves to the Queue when they have a question. Since office hours are decentralized, the Queue software is necessary to match a student seeking help with a member of the course staff. Due to this, we know nearly all office hour interactions for CS 225 were captured in the Queue logs.

3. Related Work:

As we've seen here at the University of Illinois, analyzing office hour data can provide useful information to both professors and students. There have been multiple studies done on this topic and we will discuss a few of them in this section. Instead of solely focusing on office hours, previous work has examined supplemental instruction (SI). SI can consist of peer tutoring, instructor office hours, review sessions, study groups, or any combination of these. One study found that "students who use SI have been shown to earn higher term and cumulative grade point averages (GPA's) as well as more timely graduation rates than their peers who do not utilize SI". [3]

Unlike prior studies, the office hour data collected for this paper was automatically generated every time a student posted a question on the Queue, and this data was then combined with grade data from an anonymized gradebook. On the other hand, previous studies used surveys to collect data. This usually meant that they only had students' rounded final letter grade, an approximation to how many times they went to office hours, and only had data on students who answered the survey. This causes a self-selection bias, which many researchers have noted could be "a potential threat to any deep understanding of the impact of the SI program". [4]

In a 2015 study, E. Wisniewski and colleagues looked into this topic while focusing on gender differences. The study included 941 survey responses, as well as information from the teacher, and grade data. Through this data, they found that "SI correlates with higher course grades, more confidence in the course material, greater material retention, higher overall GPA, and greater student retention and graduation rates". [3] They mostly looked to compare male and female students throughout the essay, as there seemed to be a difference between the two genders. They found that males are less likely to ask for help when needed and that female students had a higher "trigger point" (the grade at which they decided to seek out SI). [3] With the use of their findings, they were able to further improve the SI in their class.

A study in 2008 by T. Bowles, A. McCoy, and S. Bates focused on if the student will graduate on time instead of the student's grade. They believe that it is more important to look at long-term outcomes, such as graduation rates, rather than just course performance. [4] They had 3,905 people fill out their surveys and they found that SI attendance increases the probability of timely graduation by approximately 11%. In order to do this they used the treatment effects model to get rid of the self-selection problem, and then found the coefficient on SI attendance.

In 2003, the International Center for Supplemental Instruction Center for Academic Development released a national data summary about supplemental instruction. Over the course of the 5 years they looked at 745 courses with a total enrollment of 61,868 students. At the end of each year, an SI representative from each school was encouraged to fill out a summary report about the SI programs. This summary would have data on total class enrollment, SI participants' mean course grade, non-SI participants' mean course grades, and other information. They found that SI participants can have improved grades on average of 0.45 grade points higher than non-SI participants. [5]

4. Results:

What percentage of the class uses the Queue?
Based on the data presented in Figure 2, the number of times a student uses the Queue grows exponentially, with 20% of students in the course asking 82% of the questions. In the present study, 759 out of 1,238 students asked at least one question and the student with the greatest number of questions had asked 197 questions in one semester.

How do the grades of students who go to office hours compare to students who do not?
Typically, students who go to office hours earn higher grades on programming assignments than those who do not, however, these grades do not translate to exam grades, suggesting that office hours are a way for students to get help debugging their code rather than helping them understand course material (Figure 3).

The distribution of grades between students who used office hours the most and the rest of the class shows that the students who did not use office hours as much is slightly more left-skewed, with a greater percentage of those students getting grades above 95%. We believe this may be due to these students having a better understanding of the material, and thus do not utilize office hours as much. This group of students also had more outliers, with a few students in this group getting grades lower than 50% which contributed to the lower final grade averages.

Among the students who used office hours the most, the mean, median, and mode of this distribution are all around 85%. Additionally, there are fewer outliers in this category and fewer of them earned an A- or better (90%).

	Number of Students Who Primarily Went on a Given Day	Average Time Spent with Course Staff (hr:min:sec)	Average Final Course Grade
Wednesday (5 days before deadline)	53 students	**13:36** ± 13:52	**84.2%** ± 10.9%
Thursday (4 days before deadline)	50 students	**13:57** ± 11:59	**87.6%** ± 8.4%
Friday (3 days before deadline)	86 students	**14:38** ± 16:06	**85.6%** ± 10.9%
Saturday (2 days before deadline)	59 students	**12:57** ± 11:57	**84.2%** ± 10.1%
Sunday (1 day before deadline)	95 students	**13:53** ± 31:40	**83.2%** ± 11.1%
Monday (Day of assignment deadline)	275 students	**13:53** ± 13:33	**82.4%** ± 12.0%

Table 1: Students who went to office hours were grouped according to the day of the week they attended office hours the most. The table shows each of these group's average final course grade.

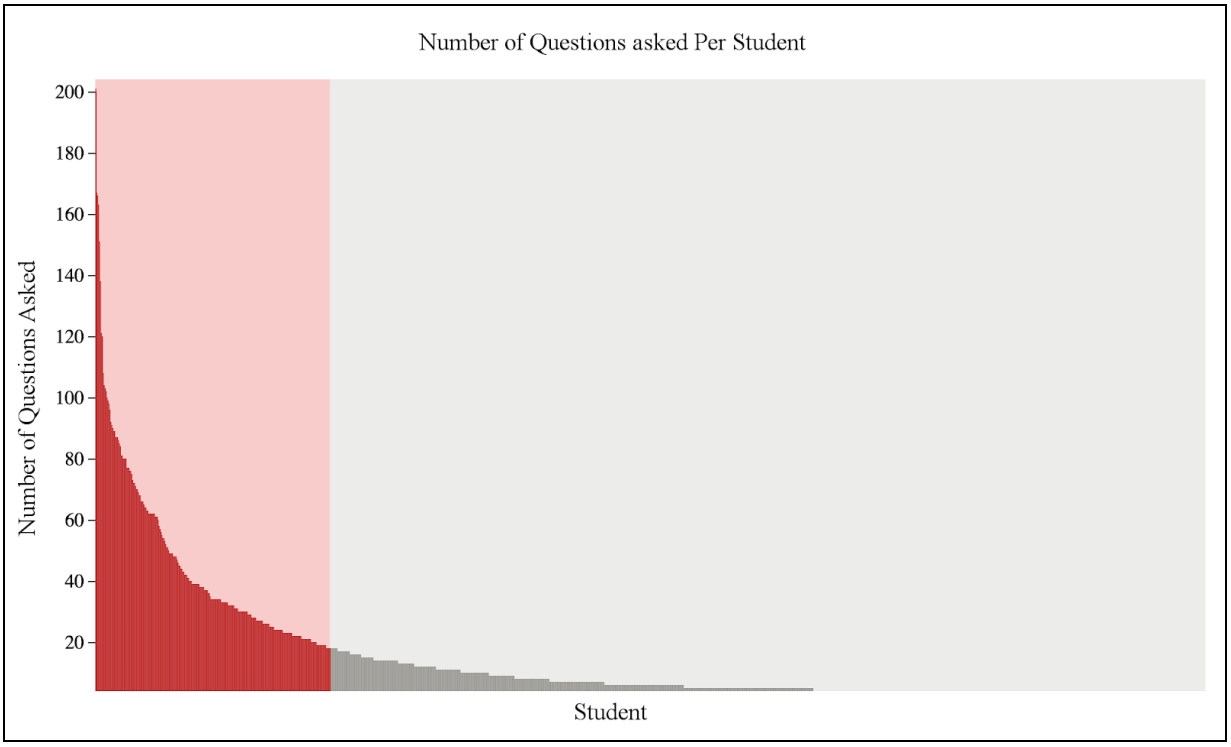

Figure 2: Distribution of the total number of questions asked by each student enrolled in CS 225 during the 2018-2019 academic year. The shaded region (left side of the graph) includes 20.03% (n=248/1,238 students) of the students who collectively asked 82.00% of the questions (n=10,225/12,752 questions).

	Top 20 % Students who Most Frequently Used Office Hours (248 students, 14-197 questions per student)	**Other 80% of Students who Rarely or Never Used Office Hours** (990 students, 0-14 questions per student)	**Significance and Effect Size** (Mann-Whitney, Cliff's Delta)
Mean grade on programming assignments	**92.13%** ± 12.20% *(Mean: 4.83% higher)*	**87.30%** ± 17.58%	p = 0.0015 δ = 0.121
Mean grade on proctored exams	**67.39%** ± 16.18%	**69.76%** ± 17.49% *(Mean: 2.37% higher)*	p = 0.0569 δ = -0.065
Mean final course grade	**84.08%** ± 9.10%	**83.24%** ± 13.80% *(Mean: <1% difference)*	p = 0.475 δ = -0.003
Normality Test Kolmogorov–Smirnov	Non-parametric Distribution KS=0.5, p=0	Non-parametric Distribution KS=1.0, p=0	

Figure 3: Average grades for the 20% of students who use the queue the most compared to the rest of the class. The ± symbol indicates standard deviation.

Does going to office hours soon after an assignment was released help the students more?
Since the programming assignments are always due on a Monday, the number of students going to office hours increases on and around this day, which can result in longer wait times. Table 1 shows that, generally, students going to office hours closer to the deadline received lower final course grades than their peers. This may suggest that students who start an assignment earlier are able to get help on less busy days, which makes the help they get more beneficial.

5. Future Work:

The present study has several limitations that prevent the generalizability of the findings reported, the most significant of which is that the results of this work are based on a single course. In order to get a better understanding of office hours and their effectiveness, a wider variety of classes is needed, preferably in a greater range of subjects. Furthermore, it would be interesting to see how the students who took advantage of office hours the most in CS 225 perform in future courses with fewer office hours.

There could have been a large number of confounding variables present in this study, which could be addressed in future studies:

Student Usage of Online Discussion Forums
Besides office hours, students are able to take advantage of online discussion forums.[1] Certain students may feel more comfortable using this online platform for getting their questions answered rather than speaking to a member of the course staff in person.

The Course Used in the Study
CS 225 has a specific setup, with most students going to office hours specifically for help debugging their programming assignments. In theory, a class where office hours are used for the sole purpose of studying for upcoming exams may help students retain course material rather than helping them debug a single assignment.

Amount of Extra Credit Given
CS 225 provides the opportunity for extra credit if students start their programming assignments early, attending lab sections, or completing daily "Problems of the Day". It is possible that certain students might feel as though they do not need as high of a score on the exams because they have a buffer of extra credit, therefore will not come to office hours for clarification on a question.

The Queue is an open source tool available on GitHub at https://github.com/illinois/queue and has been deployed at the University of Illinois (UIUC) and the University of British Columbia (UBC). If you are interested in using the queue at your institution, our GitHub page has instructions and contact information to help you and your team get started.

[1] During the full study period CS 225 utilized the Piazza online discussion forum and had a total of 5,742 questions asked on Piazza.

References

[1] Jensen, K., & Amos, J. R., & Angrave, L., & Flanagan, K., & Mussulman, D., & Schmitz, C. D., & Fagen-Ulmschneider, W. (2019, June), Adoption of an Online Queue App for Higher Education: A Case Study Paper presented at 2019 ASEE Annual Conference & Exposition , Tampa, Florida. https://peer.asee.org/32042

[2] A. Ultsch, "Proof of Pareto's 80/20 Law and Precise Limits for ABC-Analysis," *Proof of Pareto's 80/20 Law and Precise Limits for ABC-Analysis*.

[3] E. Wisniewski, R. Shapiro, E. Kaeli, K. Coletti, P. Dimilla, and R. Reisberg, "The Impact of Supplemental Instruction on the Performance of Male and Female Engineers in a Freshman Chemistry Course," *2015 ASEE Annual Conference and Exposition Proceedings*.

[4] T. Bowles, A. McCoy, and S. Bates, "The effect of supplemental instruction on timely graduation," The effect of supplemental instruction on timely graduation.

[5] "Supplemental Instruction National Data Summary, 1998 - 2003 ."