Equity-Guided Tutoring: A Scalable, TA-Run Approach

Daniel Patterson
dbp@dbpmail.net
Northeastern University
Boston, MA, USA

Josh Torre
Northeastern University
Boston, MA, USA
torre.j@northeastern.edu

Zoey Guo
Northeastern University
Boston, MA, USA
guo.zoe@northeastern.edu

Thomas McBride
Northeastern University
Boston, MA, USA
mcbride.t@northeastern.edu

Abstract

We report on a prototype small tutoring program run entirely within the existing undergraduate teaching assistant (TA) staff of a large (around 600 students) first semester Computer Science class with an equally large (around 80) staff of undergraduate TAs. The tutoring team, led by the second, third, and fourth authors (all undergraduate TAs), identified students who could benefit from the service by analyzing existing course assessments (homeworks, labs, and exams), some of which were slightly modified to aid this task.

The service was intentionally not advertised and not available by student request, so that it would not be dominated by pro-active students who were more likely to be served by the extensive traditional office hours offered by the course. The tutoring leads then matched a team of around 15 TAs with students identified to be at risk of failure and, over the course of the semester, conducted just over ninety 1-on-1 sessions with the matched students. Of the students who were tutored, the majority passed the class, and additionally, their exams scores after tutoring improved by significantly more than the student body as a whole. Additionally, qualitative feedback demonstrate the positive impact that the program had on these students.

CCS Concepts

• **Social and professional topics → Computer science education**.

Keywords

equity, tutoring, teaching assistants

ACM Reference Format:

Daniel Patterson, Josh Torre, Zoey Guo, and Thomas McBride. 2018. Equity-Guided Tutoring: A Scalable, TA-Run Approach. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

1 Background

In large introductory classes, a certain percentage of students do not pass every semester [2]. While the historic pass rates at our institution have been at or above the average worldwide, any double-digit failure to retain raises the question if there are interventions that could change these numbers. While there will always be a certain number of students who do not pass a class, one hypothesis that drove the prototype program outlined in this report is that some of those students fell *slightly* behind at some point and, due to the cumulative nature of the class, were never able to catch up. This "left behind" hypothesis posited that a limited amount of extra help, specifically targeted at students that may not even know they are in the position we identify them to be, could end up making a significant difference. While the overall number of students reached in this prototype program was relatively small—only a few dozen—the time commitment was equally small: less than one hundred hours of 1-on-1 tutoring time, out of a total of over ten thousand hours spent by the undergraduate TA team over the semester.

2 Related Work

While undergraduate teaching assistant programs are extensive, and many describe what they offer as tutoring, in the literature [4], most of these "tutoring" programs operate on a first-come-first-serve basis, at fixed hours. We consider this to be "office hours", and contrast it to our approach in §4.1. While we certainly think these are important, we believe it orthogonal to what this experience report addresses.

More related to our program, where students are proactively identified, Akhmetov et al. [1] describes a "Lost student support" program implemented as part of a larger organization structure in their large introductory courses. Unlike our program, the assistance for identified "lost students" is limited to a single 15-30 minute meeting, so it seems that they are primarily targeting students who have a small logistical challenge, rather than our program, which aims to rectify content that a student has difficulty with (and, with our program offering three to four weekly sessions, each one hour, we provide between significantly more time with the student).

Another type of program are mentoring programs that are not tied to particular courses. While these certainly have a positive impact (e.g., [3]), because they exist outside of courses, they are likely to intervene later – i.e., possibly in response to poor grades in courses – and thus require significantly more time and work (for both student and mentor) to help students get back on track. Our

program, on the other hand, requires very little additional overhead, since it is a redirection of a small subset of existing course resources, and intervenes extremely early.

3 Intervention

At the beginning of the semester, a team of undergraduate teaching assistants (TAs) was created, led by the second, third, and fourth authors. They were tasked with: A. identifying students who were at risk of failure, and thus likely currently not being adequately served by the existing course resources (lectures, labs, office hours) and B. offering 1-on-1 tutoring sessions to those individuals identified.

One key principle was that this service **would not be advertised**. It does not show up on our course website (unlike our extensive office hour schedule), and whether a student *asks* for tutoring (if they do find out about it) is not a factor in determining whether tutoring is assigned.

This is based on the assumption that if we did decide based student interest, we would end up helping students who were proactively trying to do as well as they could in the class, but not necessarily those that *needed* the help most. While obviously we do not want to discourage students seeking extra help, in order to maximize our impact, we wanted to ensure that we, the instructional staff, were deciding who to help, based on relatively objective measures of their performance relative to other students in the course.

3.1 Summary of Program

Once identified, students would then be offered 1-on-1 tutoring sessions with one of the TAs on the tutoring team. The second key principle of the service was that the **help would be limited**. As a result, the sessions were one hour long, offered weekly and generally limited to 3-4 sessions total. This time period was chosen to allow enough continuity to not only help with concrete material but help establish strategies for approaching the course work. We also wanted there to be a way for our tutors to see if those strategies were working and revise them if not, which a single, or even two, sessions would not likely support.

At the same time, the tutoring sessions were always intended to end, to allow TAs to move on to different students and ensure that eventually the students were succeeding based on their own engagement with the course material and the rest of the course resources.

The third key principle of the program was that **tutees would be chosen to maximize impact**. This led to two phases of the program, with two different identification criteria. At the beginning of the semester, we chose students who were struggling the most in class, based on concrete assessment. For example, we designed a Lab a few weeks into the semester that was review material carried out solo (unlike most Labs, which were done in pairs), and most students went through it very quickly: those that struggled with it were referred directly to the tutoring team.

Towards the middle to end of the semester, however, the team focus changed to students who had performed well at the beginning of the semester but whose performance had dropped significantly towards the middle. These were students who were at risk of failure if their performance continued on the same trajectory, but they were not the students who were doing the worst in class. Indeed, by halfway through the semester, if a student had performed poorly on nearly all the content in the course, the ability to change their trajectory in 3-4 hours of total 1-on-1 help by an undergraduate TA seemed extremely unlikely.

Students were "off-boarded" from the tutoring program based on success on one of the midterms (three were spread evenly throughout the semester), based on their own confidence with the material, or based on using up the four sessions available.

3.2 Identifying Students

We used three primary mechanisms to identify students:

- Course staff could flag someone for the tutoring team. Especially at the very beginning of the course, before we had many course assessments to use, this was used to identify students that didn't seem to be keeping up with their peers. Instructors and TAs mostly did this from Labs where they observed students directly.
- We designed two labs ("calibration labs") run in the third and seventh weeks of class, which were intentionally a review of the previous weeks of material, and intended to be solvable in less than half the allocated lab time for most students. These labs, unlike all the rest, were done solo, and the students who were still working on them at the end of the lab period (which were few in number) were directly referred for tutoring. In addition to serving the tutoring team, the labs also served as review for upcoming exams (the first and second midterms) which occurred shortly after.
- Our first midterm occurred around one month into class, and anyone who did not have a passing grade on the exam was referred to the tutoring team. Interestingly, the first midterm was also our first real opportunity to see that this program was working, as many of the students identified early in the semester passed the exam, sometimes with very good grades!
- The second midterm, which occurred around two months into the class, was used both as a way to measure success of the program and off-board students, but also as a way to identify new potential tutees. Those identified were students who performed significantly worse on the second midterm than the first, indicating that they had fallen behind relatively recently.

3.3 Working With Students

Tutoring sessions were typically scheduled on a weekly basis, with each student allowed up to four sessions per semester. This limit could be extended to five sessions if the tutor deemed necessary. Each session lasted one hour and focused on addressing the student's challenges with concepts from previous lectures. Activities included reviewing lecture notes, solving practice problems, and creating examples to demonstrate the application of specific concepts and tools taught in class. Rather than providing a similar function to office hours (where students mainly ask for help on assignments), tutoring sessions solely focused on solidifying the students' general understanding of the course material.

For sessions held shortly before an exam, the focus often shifted to working through practice exams and discussing strategies for tackling potential problems. After each session, both the TA and the tutee were required to complete feedback forms related to the tutoring program. Tutees provided feedback on the session's effectiveness, shared any complaints or suggestions, and indicated their interest in future sessions. TAs evaluated the tutee's progress and documented the topics covered during the session.

3.4 Off-boarding

Another principle we decided at the beginning was that tutoring would be a limited service. This was partly due to our own capacity: with 15 TAs adding this to their other tasks we could only dedicate so many hours per week to this, which on the scale of a 600 person class is quite modest. But it was also due to the idea that this intervention was only intended to bring a person back to where they could use the regular resources of the class: e.g., we had traditional office hours that run all day every day, with usually at least a couple TAs on at a time.

After a few weeks of running the program, we decided that our limit would be 4 ± 1 sessions, which seemed like a good tradeoff between giving students a solid amount of help while still allowing tutoring team members to move on to new students that were identified.

In addition to off-boarding based on the number of sessions, we also off-boarded students if they did well on a midterm, or if they decided they did not need the service anymore.

3.5 Benefits of Program

- Since TAs are a part of the course staff, they understand the material, pace of course, assessment, etc, very well.
- In addition to understanding the course material, TAs also have the benefit of already existing on course payrolls. Our prototype involved less than a quarter of the TAs, and each only spent an hour or two per week on this program, while still making a noticeable impact. This meant the program was extremely easy logistically to run.
- By being run almost entirely by TAs, the program easily scales up or down. The only management that the first author (lead instructor for the course) did, after assigning the second, third, and fourth authors to run the team, was to consult on some of the details about assigning work, off-boarding students, identifying students, and constructing the "calibration labs".

3.6 Challenges

The main challenge we had was ensuring that TAs who volunteered to be part of the tutoring team took on the tutees that they were assigned to. Given this was a new program, none of them had experienced it, and we additionally did not have existing training materials for them. While the skills they used in traditional office hours mostly translated, some of the logistics of communication and scheduling proved challenging.

Initially, team leads assigned tutors a student to schedule tutoring sessions with. This turned out to be less than ideal, as it was often hard to tell exactly who was meeting with students, and whether lack of meeting was because of the student never responding or the TA failed to follow up.

Of the fifteen TAs who volunteered to be on the team, two never held a single tutoring session, and of the remainder, a minority help the majority of sessions. Eventually, we switched the system to have one of the team leads (themselves TAs) managing each TA, and had the team lead handle primary communication to set up the sessions. This made it much easier to figure out when sessions occurred and ensured tutor accountability. This change was also positively received by tutors who, when asked what could be done to make the program better, many responses included the continuation of this style of delegation.

4 Alternatives

There are two main alternatives for students, currently: office hours (offered in all classes) and external "peer" tutoring, currently most commonly through a program called Knack. It is also possible for students to find paid tutoring services, but it is likely they suffer the same issues as external peer tutoring, so we don't address them separately.

Each has significant drawbacks:

4.1 Office Hours

While these are the main resource offered by the course, and are incredibly important, they intentionally are equally available to all students, regardless of how well they are doing in the class. While we can expect students that are at the very top of the course may not use office hours much (or at all), there is a spectrum of students who use them, and there is no easy way for TAs holding office hours to allocate help to students based on need, rather than equally.

There are two primary reasons for this. First, office hours are often carried out in large shared rooms, with public queues identifying how many students are waiting. While at certain times of day or days of the week they may be quiet enough to allow significant 1-on-1 help, at others, TAs must move through students steadily to not draw the ire of others in the queue.

Secondly, until a TA is interacting with a student, there is no way for them to know how much help they need, and so if they were to try to allocate help based on need, they would have an impossible task: perhaps they spend half an hour helping one student, who seems like they need a lot of help, only to have the next student need even more.

While it is certainly possible to attempt to establish a form of triage for office hours, and sometimes this is done in order to find common questions, it likely would not work to allocate help based on need. The reason for that, of course, is that all students who are struggling, which attending office hours is evidence for, deserve to be helped!

Indeed the point of the program described in this paper is not that we disagree with the idea of helping all students, but rather that we see a potential for bringing students from a point where resources like office hours are insufficient to help them to a point where they are sufficient to help them. Clearly, such a program is not necessary for someone for whom office hours are already working.

4.2 External-to-course peer tutoring

The other resource that is available to students are so-called "peer tutoring" systems, where students can sign up to tutor particular subjects 1-on-1, and then students in such courses can sign up to be tutored.

The problem with these services are: A. the tutors are only vetted by the external service, not by the instructional staff, and so may vary widely in their understanding of the course material, even for courses that do not change much year-to-year. For courses that do change, the tutor may have to learn the material as they are tutoring, diminishing their capacity to help. B. by being external, they have no way of coordinating with other resources, asking for advice from course staff, passing students on to the instructor, suggesting interventions in lab, etc.

4.3 An Internal Tutoring Team

Both of these issues are resolved by having a dedicated internal tutoring team. By identifying and only working with students who have demonstrated the most need, 1-on-1 time, outside of office hours, can be set up without worry about not allocating time most efficiently, or being "unfair" during office hours. And since the team is deeply integrated into the course staff, and composed of TAs that are known to understand both the course material in general and the specifics of what is happening in the course, in terms of assignments, labs, etc – they do not suffer from the lack of utility that sometimes peer tutoring runs afoul of.

5 Assessment

We present this in two sections: first, we present quantitative data based on how students who we enrolled in the tutoring program did both in the course as a whole and on exams, both before (and compared to the course as a whole) and after.

Next, we present qualitative assessments of the program, drawn from post-tutoring-session surveys that TAs and students completed, an end of semester survey, and from unsolicited comments sent to tutors. While the program was small in scope, we hope that this limited analysis can be useful in showing that it did work as designed, and help others who wish to create similar programs or replicate our design.

5.1 Quantitative

Over the semester, tutoring sessions were administered to a total of 43 students who held a recorded 92 sessions (although this number may be higher as, especially at the beginning, some TAs may have forgotten to report the session).

There are three distinct groups of students we will analyze, referred to as the Pre-Exam, Exam 1 and Exam 2 groups. The course had three mid-term exams and no final: Exam 1 occurred one month into the course, Exam 2 occurred two months into the course, and Exam 3 occurred three months into the course (which was very close to the end of the course, though before the course "drop" deadline).

The Pre-Exam group is made up of students who were identified by the "benchmark" lab described in §3.2, or who were referred directly based on observation by course staff (in lab or elsewhere). The Exam 1 group is made up of students who failed the first exam.

	Pre-Exam	Exam 1	Exam 2
Total Students	12	17	14
Students that Dropped	5	5	3
Student Drop Rate	41.67%	29.41%	21.43%

Table 1: Drop Rate by Tutoring Cohort

The Exam 2 group we took a hybrid approach. After filtering for students who did well on the first exam, we began with the students who scored poorly (45%) on the second exam and continued assigning tutors based on capacity, reaching eventually to scores in the mid 70s.

The inaugural goal of this program was to reduce the chances of course failure for students who were identified as "at risk". Since students know all of their exam grades before the drop deadline, this translates to the drop rate acting as an effective failure rate for students. Table 1 shows the drop rates for each of the aforementioned cohorts of students.

While the rates for the Pre-Exam and Exam 1 groups are quite high, we note that these students were all chosen because we believed that without intervention, these students would all fail. Following this hypothesis, our intervention resulted in a majority passing when they may not have otherwise. The Exam 2 group was different: while we were still certainly worried that the downward trajectory between the first and second exam would continue to the third, and thus they may have risked failure, some of them may have been able to scrape by with a passing grade. This generally stronger performance is reflected in the increased pass rate of that cohort.

We also used exam scores to compare student progress in the course, as the exams are cumulative in nature and are held in a controlled environment. This data, shown in Table 2, shows a significant correlation between the administration of tutoring to students and their success in subsequent assessments. In particular, students in both the Exam 1 Cohort and Exam 2 Cohort make a significant increase on the exam following their sessions (Exams 2 and 3 respectively), which is controlled against the class as a whole, indicating that this was not because the subsequent exams were easier. Indeed, the data indicates that Exam 3 was likely slightly more challenging than Exam 2. We do not include data on the Pre-Exam group because we have no objective measure of them before the tutoring intervention.

Additionally, in data collected in a required exit survey, students rated tutor efficacy an average of 9.45 out of 10. While clearly subjective, this leads into our next section – that tutoring can help students feel valued by the course staff.

5.2 Qualitative

While the above quantitative data shows the program had a positive impact, this data alone does not completely represent the impact tutoring has on students who are learning Computer Science. There are increased societal and economic pressures in the current state of the industry to be the "best" which leads to rampant imposter syndrome [5, 6]. This also may discourage students to stay in the field.

	Exam 1 → 2 (Exam 1 Cohort)	Exam 1 → 2 (Entire Class)	Exam 2 → 3 (Exam 2 Cohort)	Exam 2 → 3 (Entire Class)
Mean Change	+10.54%	+0.29%	+15.14%	-3.45%
Median Change	+18.79%	0.00%	+13.89%	-2.44% %

Table 2: Average Change in Exam Score After Tutoring vs. Control

Tutoring allows for students to increase their confidence in the course, especially in introductory courses where there are a breadth of skill levels and personal experiences. Additionally, the existence of the tutoring program as an official part of the course, rather than an external service, helps to normalize it.

Students who responded to an end-of-semester form said: "I felt [tutoring] was a good way to get a one on one understanding of the material especially since a large lecture can be hard to focus", "Tutoring helped me understand and piece together the big topic[s] of [our intro CS course] and clear up any confusions I had", among other positive personal anecdotes. In one of the post-session exit surveys, a tutee captured the positive impact of the pro-active design of the program: "[The tutor] reached out to me and helped me without me seeking anyone. This was the best scenario I could've asked for."

Tutees also often sent unsolicited positive messages to tutors, especially after exams, e.g., "Hey [tutor name] the exam results r out n holy this is my highest score I just wanted to thank you, u r a wizard", and even in cases where they did not perform as well as they wanted to, the individual attention still seemed to help students, e.g., "I got #destroyed by this exam but I wanted to take the time to thank you properly big bro. I know we only had a few sessions together but it really meant the world to me being able to actually get help in that class one on one. Thank you so much for your time this semester I hope you enjoy your break [tutee name]!"

6 Conclusion

While we have run this only once, we believe that this structure could easily be replicated in other courses, and the evidence, while limited, seems to suggest that it would be beneficial to do so. The administrative overhead at the instructor level was very low, as it was almost entirely run by undergraduate TAs (the second, third, and fourth author of this report). The total time consumed by all the TAs involved in the program was also, comparatively, quite low: with a staff of 80, we had a hypothetical 1200 hours per week, of which we were only dedicating a maximum of 20 or so towards tutoring, or 5%. To get the results we got for 5% of TA time seems absolutely worth it, at least in introductory classes.

Acknowledgments

We thank all the members of the tutoring team for time, effort, and care they put into helping students they were working with.

References

[1] Ildar Akhmetov, Sadaf Ahmed, and Kezziah Ayuno. 2024. How We Manage an Army of Teaching Assistants: Experience Report on Scaling a CS1 Course. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1* (Portland, OR, USA) *(SIGCSE 2024)*. Association for Computing Machinery, New York, NY, USA, 32–38. https://doi.org/10.1145/3626252.3630871

[2] Jens Bennedsen and Michael E. Caspersen. 2019. Failure rates in introductory programming: 12 years later. *ACM Inroads* 10, 2 (April 2019), 30–36. https://doi.org/10.1145/3324888

[3] Shearon Brown and Xiaohong Yuan. 2014. Experiences with retaining computer science students. *Journal of Computing Sciences in Colleges* 29, 5 (2014), 34–41.

[4] Diba Mirza, Phillip T. Conrad, Christian Lloyd, Ziad Matni, and Arthur Gatin. 2019. Undergraduate Teaching Assistants in Computer Science: A Systematic Literature Review. In *Proceedings of the 2019 ACM Conference on International Computing Education Research* (Toronto ON, Canada) *(ICER '19)*. Association for Computing Machinery, New York, NY, USA, 31–40. https://doi.org/10.1145/3291279.3339422

[5] Adam Rosenstein, Aishma Raghu, and Leo Porter. 2020. Identifying the Prevalence of the Impostor Phenomenon Among Computer Science Students. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (Portland, OR, USA) *(SIGCSE '20)*. Association for Computing Machinery, New York, NY, USA, 30–36. https://doi.org/10.1145/3328778.3366815

[6] Angela Zavaleta Bernuy, Anna Ly, Brian Harrington, Michael Liut, Andrew Petersen, Sadia Sharmin, and Lisa Zhang. 2022. Additional Evidence for the Prevalence of the Impostor Phenomenon in Computing. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education - Volume 1* (Providence, RI, USA) *(SIGCSE 2022)*. Association for Computing Machinery, New York, NY, USA, 654–660. https://doi.org/10.1145/3478431.3499282