Who Benefits? Positive Learner Outcomes from Behavioral Analytics of Online Lecture Video Viewing using ClassTranscribe.

Lawrence Angrave
Zhilin Zhang
Genevieve Henricks-Lepp
Chirantan Mahipal
angrave@illinois.edu
zhilinz2@illinois.edu
henrick2@illinois.edu
mahipal2@illinois.edu
University of Illinois at Champaign-Urbana
Urbana, Illinois

ABSTRACT

PREPRINT. Not for publication or general distribution. Only for University of Illinois internal use.

Lecture material of a sophomore large-enrollment (N=271) system programming 15-week class was delivered solely online using a new video-based web platform. The platform provided accurate accessible transcriptions and captioning plus a custom text-searchable interface to rapidly find relevant video moments from the entire course. The system logged student searching and viewing behaviors as fine-grained web browser interaction events including full-screen-switching, loss-of-focus, incremental searching events, and continued-video-watching events with the latter at 15-second granularity. Student learning behaviors and findings from three research questions are presented using individual-level performance and interaction data. Firstly, we report on learning outcomes from alternative learning paths that arise from the course's application of Universal Design for Learning principles. Secondly, final exam performance was equal or better to prior semesters that utilized traditional in-person live lectures. Thirdly, learning outcomes of low and high performing students were analyzed independently by grouping students into four quartiles based on their non-final-exam course performance of programming assignments and quizzes. We introduce and justify an empirically-defined qualification threshold for sufficient video minutes viewed for each group. In all quartiles, students who watched an above-threshold of video minutes improved their in-group final exam performance (ranging from +6% to +14%) with the largest gain for the lowest-performing quartile. The improvement was similar in magnitude for all groups when expressed as a fraction of unrewarded final exam points. Overall, the study presents and evaluates how learner use of online video using ClassTranscribe predicts course performance and positive learning outcomes.

CCS CONCEPTS

• **Information systems** → **Video search**; *Crowdsourcing*; *Speech / audio search*; • **Human-centered computing** → **Accessibility**; • **Applied computing** → **Education**; • **Computing methodologies** → *Speech recognition.*

KEYWORDS

learning,accessibility,captions,video-search,transcription-search, learning-analytics,behavioral-analytics,student-behavior

ACM Reference Format:

Lawrence Angrave, Zhilin Zhang, Genevieve Henricks-Lepp, and Chirantan Mahipal. ACMYEAR-TBD. Who Benefits? Positive Learner Outcomes from Behavioral Analytics of Online Lecture Video Viewing using ClassTranscribe.. In *SIGCSE '20: ACM Symposium on Computer Science Education,March 11–14, 2020, Portland, OR.* ACM, New York, NY, USA, 7 pages. https://doi.org/ACMDOI-TBD

1 INTRODUCTION

The primary goal of this paper is to report on and evaluate the replacement of physical lectures (that used a blend of active learning and instructor-led live-coding examples) with online content designed for online viewing. These videos were delivered using a new custom web-based, text-searchable video player system, ClassTranscribe, that met accessibility standards, applied Universal Design for Learning principles, and allowed students to search for relevant content by indexing transcribed video. The ClassTranscribe application generated low-cost accurate captions by combining automated and crowd-sourced techniques. The latter included editing by both course staff and students within the course. Learning outcomes of the Spring 2019 students are compared to those of previous comparable semesters. By combining behavioral interaction data captured by ClassTranscribe with student "gradebook" assessment data, the learning benefits of the tool and online videos were examined by class quartile within the course and also compared to three previous comparable semesters.

SIGCSE '20, March 11–14, 2018, Portland, OR

https://doi.org/ACMDOI-TBD

In Spring 2019, Monday-Wednesday-Friday lectures for a sophomore system programming class for CS majors were scheduled at 8am. The instructor suggested this early time would severely impact attendance and learning. Rather than penalizing absence, the instructor offered to record equivalent lecture content specifically designed for online viewing. The scheduled time slots were used three times: once for the first lecture, once for a guest lecture, and once for the closing lecture. The weekly Wednesday lecture time was re-purposed for instructor office hours, which were utilized by two students over the semester, and the the Monday and Friday meeting times were not used. This change was announced only on the first day of class. The instructor offered to provide in-person 8 am lectures for students who requested an in-class experience, but no student requested it. Previous semesters also offered online video lectures, but the content consisted of simple automated recordings of the classroom lectures. For this semester, only the new recorded video lectures were used by the course and delivered by the tool presented in this paper.

This paper is organized as follows. In the background section, we review the use of lecture videos and Universal Design for Learning (UDL) principles from both the CS education and educational literature, and we briefly discuss students' use of videos and searching for content as a means of regulating their metacognitive awareness. In the related work section, we introduce the pedagogical utility of the new tool by enumerating its unique features that are only sparsely represented in other video delivery systems. In the technology aspects section, we describe interesting technical details from the tool development activities and platform IT requirements. In the learning outcomes section, we evaluate the effect on student learning outcomes and utilization of UDL resources. Limitations and opportunities for further work are discussed in context throughout the paper.

2 BACKGROUND

Online learning environments constitute an important application area of human-computer interaction and are ubiquitous in today's educational landscape. A key instructional element in these learning environments is the set of video lectures that often serve as the unique portal connecting the instructor to the learner, apart from serving as primary portals of content delivery [18, 20, 23]. Not surprisingly, studies have found that it is the course component with which students spend most of their time [3, 29]. Other studies, especially those on massively open online courses (MOOCs), have found that a significant number of students primarily watch only videos, while skipping over assessment problems and other interactive course components [1, 16]. Students report that the videos make them feel more engaged in the course and motivate them[11], and that they also feel more connected to the instructor [14, 26]. Additionally, students perceive that the videos help them learn [27]. However evaluating if the videos have actual effects on learning is a necessary criteria for most practitioners, and this paper investigates video watching beyond affective measures and perceptions to better understand learning outcomes, specifically in videos that are closed-captioned and transcribed.

Videos used in the study were closed-captioned and had accompanying transcriptions, which adhere to Universal Design for Learning (UDL) principles. Following these principles ensures that students' diverse abilities and needs are met. A UDL framework, [5] suggests that instructional materials encourage (1) equitable use, (2) flexibility in use, (3) simple and intuitive use, (4) perceptible information, (5) tolerance for error, (6) low physical effort, and (7) adequate size and space for use. By providing materials that incorporate these principles, students of all backgrounds have fewer barriers to learning [25]. Although research has found that students enjoy courses that adhere to UDL principles (e.g., [22]) and that they perceive their instructor's teaching abilities to be better[9], there do not seem to be significant differences in learning outcomes for students with and without disabilities. For example, [15] found no significant difference in both disabled and non-disabled students' learning between a course taught using UDL principles compared to a course taught in a regular manner. [10] found improvements in learning outcomes for some types of UDL materials but not all. This study includes the UDL materials, in the form of videos accompanied with text, and secondary learning opportunities of the online course book, to understand if this material promotes better learning outcomes for all, and if alternative learners paths that utilize these materials result in positive learner outcomes.

One reason that we hypothesize that videos adhering to UDL principles–and videos that have a text-searchable feature in particular–will be beneficial for all students is because of their ability to assist in students' metacognitive awareness. Metacognition is awareness of one's own thinking; more specifically, it enables a person to monitor, assess, and regulate one's understanding and thought process [12]. Thus students who are metacognitively aware are able to recognize what they do not know and then figure out how to extend their knowledge [2]. As such, engaging in metacognitive thinking and strategies improves learning outcomes [19, 21, 34], and this effect on learning outcomes also holds true in online environments [4] . Engaging in metacognitive strategies is effective for improving low-achievers' learning outcomes [6], and it also explains the difference between high and low achievers[17, 33]. Instructors must be deliberate in their support of metacognitive awareness [28]. In this particular course, students had access to a unique text-searchable interface to rapidly find relevant video moments from the entire course. Thus, they were able to reflect on what they did not understand and could easily locate material to assist them in clarifying that understanding. [32] developed a similar tool, although no research to date has looked at the relationship between the tool's use and student performance–a gap in which this study attempts to fill. In the next section we outline related tools and discuss the utility of the tool presented here.

2.1 Related Work

Modern creation and delivery of educational online lecture videos includes a multitude of presentation formats, sources, and delivery tools. Sources include lecture-room multi-view recording systems to simultaneously capture live lectures from room camera and display sources, screen recording applications (e.g. Camtasia, OBS Studio and native support in Microsoft Windows 10), and content professionally edited for MOOCS and large audiences. Delivery mechanisms span emailing a hyperlink to the class ("Here's the mp4 file on my shared directory") to sophisticated commercial systems that optimize video playback for large audiences (e.g. Youtube, echo360, Kaltura, Coursera).

Recent work on crowdsourcing transcriptions include [7, 24, 32]. The system presented here overcomes limitations from existing open-source and commercial alternatives. Namely, it is a system that enables all of the following features:

- Has inexpensive and accurate captions. Commercial caption systems offered free-but-low-accuracy transcriptions, or professional transcriptions services at $1 per video-minute.
- Has fine-grained server-side logging of user interaction events to support educational research, student engagement and course rewards based on students exhibiting desired behaviors. The former includes both insights into how and when students are engaging with course videos resources and, by joining with course performance data, provides insights into how different learning behaviors benefit each student or student sub-populations.
- Deploys easily on University hardware or the commercial cloud.
- Is open source and extendable. The complete source-code of the web application is available at https://github.com/cs-education/classTranscribe.
- Supports modern accessibility web standards (e.g. closed-captions with configurable display settings, Aria-tag support for blind users)
- Supports multi-stream and viewer-controlled playback speed.
- Enables the finding of relevant content by indexing transcriptions within the current video and across the entire course that is relevant to the student's current activity or the student's pre-exam review knowledge-seeking question.

In the tool presented here, audio is initially transcribed using a modern high-accuracy speech-to-text cloud service (Azure Cognitive Services Speech-To-Text) at a cost of approximately $1 per hour. This service also supports improved transcription accuracy of domain-specific speech (though this feature was not enabled or evaluated in this study, and is an opportunity for future tool improvements and research).

A similar comparable system is the Indexed Captioned Searchable Videos ("ICS") system used at the University of Houston [32] and [31]. Recent evaluation of ICS in an introductory psychology course [30] reported increased instructor and student satisfaction but "confounding variables in the structuring of the course prevent us from making firm conclusions about student performance."

3 TECHNOLOGY ASPECTS

An early version of the software that included automated speech to text was first available in November 2018. Course videos were automatically enumerated and downloaded by polling playlists from an existing commercial lecture capture system. However this paper reports on a Spring 2019 system programming course for CS majors where ClassTranscribe was used exclusively to deliver lecture content designed for online viewing. The lecture videos closely mirrored the lecture exposition and problem-solving lecture material used in previous semesters.

The Web accessibility of ClassTranscribe was evaluated using the AInspector WCAG Firefox plugin [13], and a University web-accessibility reviewer also manually reviewed it. Specific issues found were entered as bug reports. Correcting accessibility errors were useful and appropriate bug-fixing activities for new programmers to a project because they had the following characteristics: i) Shallow, front-end errors that were usually specific to a specific page or single resource; ii) Easy to find, reproduce, test and fix iii) Motivating and small iv) Had a high perceived value to new developers v) Encouraged developers to appreciate accessibility concerns from the earliest involvement with the project.

The system programming lecture videos were recorded using the free, cross-platform OBS Studio software using the H.264 MPEG encoder at "High quality Medium Filesize" application settings. OBS Studio was configured with two recording layouts that could be switched while recording: Head-shot from the inbuilt laptop camera, or laptop full-screen with a small "postage stamp" head-shot inlay. The former was only used for introductory and closing remarks. A traditional 50-minute lecture was recorded as a series of small problem-centered video segments (typically 4 to 10 minutes). No post-editing was performed, as the instructor believed it was quicker to re-record the segment than fix errors. The total semester content for the course compromised of 37.0 video hours and 5.2 GB of 286 mp4 files. Bandwidth requirements of introductory full-screen talking head format were 200 KB/s, however the bulk of the content — screen annotations and sketches (created using ScreenBrush.app for OSX and a Wacom tablet pen-based input device) and developing program code using a text editor and terminal shell — compressed to an average 40 KB/s.

The technology of the ClassTranscribe system was comprised of common HTML 5 components (Bootstrap and jQuery), Node.js web server, and a SQL database. The system was built and deployed as a set of Docker containers on a dedicated Linux virtual machine. The database schema and design choices were published in [7].

The video server was deployed to a single virtual machine hosted on-campus on a DL380 Gen 8 2014 server with 1 Gbit/s network connectivity, 12 CPU cores, 15 GB RAM, 500 GB SSD storage, with the expectation that this would be sufficient to support at least 500 concurrent users. In practice, due to the course's small video file size and bandwidth (typically 10 - 20 MB and 40 KB/s respectively), neither network, storage, or CPU needs of a single host machine were limiting factors. Content for other courses (including a computer science theory course, statistics, and machine learning course) from lecture-capture sources had larger files and bandwidth requirements,> 200 MB, 4 MB/s respectively, but comprised only 5% of the total video minutes viewed. A future version of the system could also pre-process external large files to be smaller, which would also benefit mobile users on low bandwidth connections.

4 LEARNING OUTCOMES

4.1 Data Collected

Event logging of student interactions was available in approximately the last 7 weeks prior to the final exam, of which 6 weeks contained lecture video content. Table 1 presents a summary of the events collected for the system programming course by event type. A *timeupdate* event was generated whenever a user watched 15

Table 1: A summary of the 1198413 events by user action.

Action	Count	Fraction
timeupdate	413912	0.345383
play	164428	0.137205
seeking	160924	0.134281
pause	153034	0.127697
userinactive	134455	0.112194
seeked	102531	0.085556
changevideo	24750	0.020652
changedspeed	23151	0.019318
fullscreenchange	17195	0.014348
selectcourse	2850	0.002378
filtertrans	1087	0.000907
edittrans	83	0.000069
sharelink	13	0.000011

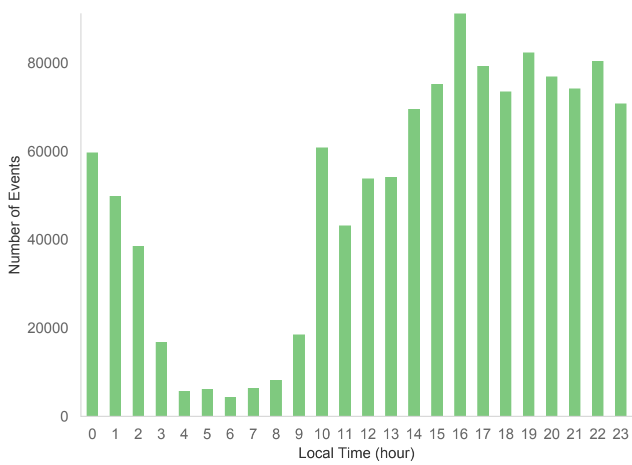

Figure 1: Events captured by hour

continuous seconds of video, which could take fewer wall-time seconds if the video was played at a higher speed. A total of 1725 video hours was served over the last 7 weeks of the course. The small number of *selectcourse* events suggest students remained logged in between viewings. The small number of times that students shared a URL link to a specific moment in the video suggests students may have been unaware that they could easily share and discuss a particular video moment in the course's discussion forum external to this tool. This suggests a future version of the tool should present user tips to enable more pedagogically valuable utilization of the tool. Students used the tool throughout the day, particularly in the afternoon and evening (See Figure-1). Only 0.6 % students chose to learn at the scheduled 8am time slot.

4.2 Evaluation

In the second part of the paper, we use behavioral and gradebook data to understand how students learn and thrive in the course. We present three student-outcome questions -

- “Are students exploiting multiple learning paths?" Are students utilizing the same learning resources in similar proportions, or are they exploiting multiple learning paths available and thriving?
- “First, do no harm" - Did the cancellation of traditional in-person lecture and sole use of the ClassTranscribe for lecture content, affect exam performance when compared to similar previous semesters?
- “Who benefits?" Do online lecture videos help all students, or only the strongest students? Is more video minutes use associated with improved exam scores?

4.3 Learning Outcomes of Multiple Learning Paths

Students may learn about a particular system programming topic by attending a lecture (2015-2018), viewing a recorded lecture from a commercial capture system (2015-2018), viewing lecture video recorded for online use with ClassTranscribe (in Spring 2019 semester only), searching and/or reading lecture transcript (Spring 2019 semester only), reading the free course book in a variety of formats (epub,pdf,html; in all semesters), or viewing recordings of classroom lectures (previous semesters only). This is an example of application of UDL principles, where multiple learning activities and resources can be utilized by the student to achieve the same learning goal. Using the behavioral data captured in Spring 2019, it became possible to compare learning outcomes of alternative learning activities and validate student use of UDL materials in the course's design.

In Spring 2019 semester, out of 271 students taking the final exam, 194 (71.6%) students used ClassTranscribe and 77 students never viewed videos in ClassTranscribe (28.4%). This group was identified as the "Course Book" learners, as this was the most likely source that the students used to learn the content. It is possible a subset of students found and viewed previous semester lecture videos on the commercial lecture capture system, or co-watched the intended

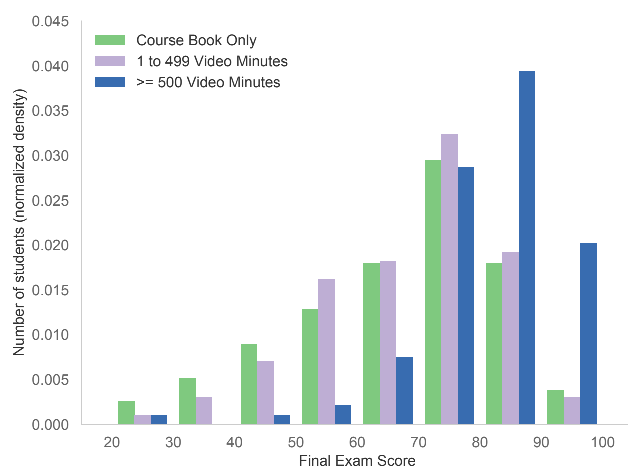

Figure 2: Histogram of final exam scores for different learning paths.

Table 2: Final exam score (out of 100) for students who chose to learn without using ClassTranscribe (Path-I), watched below average duration of video content in ClassTranscribe (Path-II), or above average duration (Path-III).

	Learner Path-I $n = 78(28.8\%)$ 0 minutes	Learner Path-II $n = 99(36.5\%)$ 1 to 499 minutes	Learner Path-III $n = 94(34.7\%)$ $\geq$ 500 minutes
Q_1	57.0	59.8	76.5
Q_2	70.4	71.4	82.4
Q_3	77.8	79.4	88.8

videos while another had logged in. The online course book is the most plausible learning path because it was positively mentioned in course feedback forms and linked from the course web pages. The 194 students who viewed at least 15 seconds of continuous video in ClassTranscribe (i.e. generated at least one *timeupdate* event) watched an average of 500.1 video minutes (equivalent to 8.3 video hours, and was rounded to 500 video minutes in subsequent analyses). Behavioral data were captured over 7 weeks, of which 6 contained content, corresponding to 83 minutes viewed per week. This time-on-task is lower than 3 live lectures total 150 minutes per week.

Students' primary choice of learning activity were characterized as 3 learning paths: i) No use of video lectures (i.e., the course book readers) ii) Below average use of video lectures iii) Above average use of video lectures. Reading events of the course book were not available. A limitation of this analysis is students may further supplement these expected learning activities by engaging in other relevant but untracked resources (for example, web searches, discussion forums, and discussions with their peers and course staff).

The final exam quartiles and histogram of the learning outcomes of the 3 learner paths are summarized in Table-2 and Figure-2 respectively. The primary finding is above average usage of ClassTranscribe led to improved exam performance. For example the median of Path-III learners was $\geq$ 11.0 point improvement on the final exam compared to learners choosing Path-I or Path-II.

A secondary finding is that a significant fraction of students can still earn competent – albeit on average lower – scores in a final exam by choosing alternative and secondary learning resources (Learner Path-I). This is encouraging feedback and validation for Universal Design advocates and practitioners who have created, or are considering creating, equivalent course resources in alternative formats.

4.4 Comparison With Prior Semesters

The final exam over all previous semesters used 45 randomized multiple choice questions (with randomized variants) that covered all topic areas of the course, and included concept questions and skilled application of a system programming idea within a programming context. Long form questions varied by semester, were graded manually using a rubric, and were excluded in this analysis. The exam scores of the Spring 2019 semester were compared to previous Spring semesters with the same instructor. Though programming and lab assignments have varied incrementally over time, the overall content, structure, lecture delivery and exam assessments have remained similar. The quartiles of multiple choice final exam scores are presented in Table-3. All quartiles of the Spring 2019 semester show a modest improvement compared to previous semesters. A counter argument is to observe that quartiles show a mostly slow rising trend, and we hypothesize that active and blended learning in physical lectures or use of lectures videos has an insignificant role in this course. In either perspective, we can still conclude that the *dropping of lectures caused no significant harm* to exam performance in a lower-curriculum required class of the CS undergraduate program.

Can we therefore continue and conclude that lectures have no other intrinsic or extrinsic additional value? No. A limitation of this finding is that we have reduced learning outcomes and lecture value to the numerical performance in a final exam setting and not measured a student's experience and affect for Computer Science e.g. long-term interest, motivation, perseverance, and the ability to thrive in future academic and non-academic settings.

Table 3: Final exam multiple choice quartiles and 50% spread ($Q_3 - Q_1$), for students in Spring semesters of 2015,2016,2018, and 2019.

Statistic	2015 $N = 195$	2016 $N = 325$	2018 $N = 247$	2019 $N = 271$
Q_1	60.1	58.7	60.0	**63.6**
Q_2 (median)	68.1	69.6	71.1	**72.7**
Q_3	76.6	78.3	80.0	**83.3**
$Q_3 - Q_1$	16.5	19.6	20.0	**19.7**

4.5 Learning Outcomes By Quartile

Lower performing students (Q_1 Q_2) may struggle to succeed due to diverse course-related, affect, and acquisition reasons (e.g. missing prerequisite knowledge or skills, incomplete, ineffective or inefficient learning strategies), as well as external non-course hardships (e.g. financial stress, limited time, reduced peer support, lower confidence, imposter syndrome). Apriori, i) It is unclear if ClassTranscribe could be used effectively by lower-performing students (compared to their peers) to improve their course-exit performance as measured by their final exam score; ii) However, the tool provided an opportunity for students to find and review relevant lecture material on any day and any hour of the day. Did this help ameliorate difficulties experienced by lower performing students? Similarly, could upper-performing students also gain from its use, compared to their peers?

The last analysis presented here explored if learning benefits were possible for both lower performing and upper performing students in the Spring 2019 semester. Students were grouped into performance quartiles based on their total weighted course score (with the final exam score excluded). The course score was comprised of numerous assessments that is typical of a CS programming-heavy course: challenging autograded weekly programming assignments, quizzes and programming problems given under exam conditions, and lab programming exercises. All assessments were individual work.

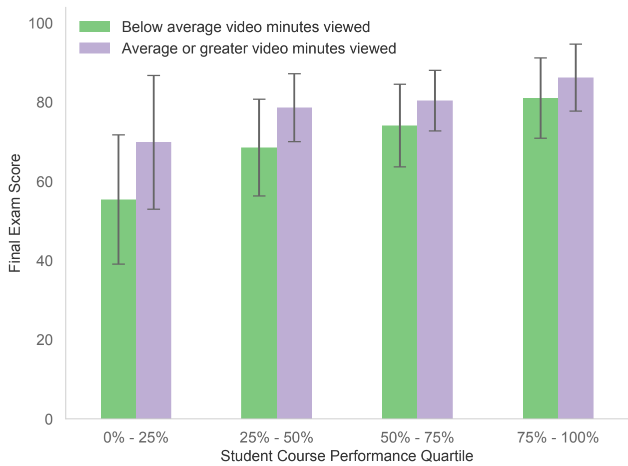

Figure 3: Overall Course Performance partitioned by Quartile Rank

The total class enrollment (N=271 exam takers) was sufficiently large to evaluate 4 quartile groups each split into 2 subgroup conditions (below vs. above threshold of total-video-minutes-viewed). To compare ClassTranscribe use within each group, the average of video-minutes-watched-by-that-group-only was an appropriate threshold because it i) is a simple general definition that can be implemented in future analyses and replication studies; ii) is statistically expedient and likely to ensure a reasonable sub-population size; iii) allows the comparison of student outcomes for students of a similar competency. Using a threshold based on total class average or fraction based on total video minutes were considered but rejected in favor of a per-quartile threshold due to the above 3 reasons.

For each quartile, students who used ClassTranscribe more than average for that quartile, surpassed their peers' exam performance in the same quartile by a significant amount (Table-4 and Figure 3)). For example, on average, in the lowest performing quartile, a student could increase their exam score by 14.4 absolute points (out of 100) compared to another student in the same quartile who did not significantly use the tool. A Mann-Whitney U test [8], which requires no normal distribution assumption, rejects the null-hypothesis. The exam performance increase for every quartile is significant at the $p < 0.05$ level.

The largest increase in exam improvement was exhibited by the lowest performing students (Q_1) using least additional studying (5.6 hours over the 7 week period). This supports the finding that use of ClassTranscribe is an effective learning technique that is open to students of all competency levels.

A second finding is that the benefit to all students is similar in range (30 – 40%) when expressed as fractional reduction of average points lost on the final exam for each quartile. Replication of this fractional effect, understanding its magnitude, cause and scope, is worthy of future research. Further, a reasonable message to students of all abilities is, "Want to do well in the final? Take time to watch the videos over the semester and you can reduce your exam points you would have lost by a third."

5 CONCLUDING REMARKS

We reported on the removal of lectures from a large enrollment system programming class and the addition of online lecture content designed for online viewing, using a custom accessible video system that supported learning analytics.

The fine-grained student behavioral data captured from the tool was used successfully to explore varied aspects of student learning. Learning outcomes both within the course and across comparable semesters provided evidence that use of video lectures using ClassTranscribe is an effective learning technique that leads to positive learning outcomes for students of all competency levels.

Student behavioral data also provided insight into actual and successful student use of alternative learning paths available due to the application of Universal Design for Learning principles in the course design.

This paper examined learning outcomes with respect to exam scores and found statistically significant effects with video minutes viewed. It is the first paper to report and evaluate learning outcomes of the ClassTranscribe. The potential for future research and education insights using these (and similar future data) is broad and significant, including for example, machine learning models and multivariate statistical models to predict student outcomes and behaviors, and Hidden Markov Models to model event sequences. The results presented were based on a single measure - video minutes viewed. Future work will explore other learner behaviors that are possible with the data presented here (for example, engagement effects due to full-screen viewing, effects of search and transcription use, and comparing video minutes to wall-clock minutes to determine which is a better predictor of learning in technically dense computer science content).

The results and experience of this study were also used to identify limitations and prioritize future improvements to the tool (for example, encouraging greater use of search and share functionality and improving accessibility).

Practitioners interested in replicating or adopting this approach may use the repository to deploy the system to their own server, which requires only modest hardware requirements, or use the existing server at classtranscribe.com and supported by email .

ACKNOWLEDGMENTS

We wish to thank the Illinois students who contributed to the ClassTranscribe project, and members of the Computer Science and Education group and the Illinois iLearn Group who helped in the preparation of this paper. This work is supported by a Microsoft Corporation gift to the University of Illinois as part of the 2019 Lighthouse Accessibility Microsoft-Illinois partnership.

REFERENCES

[1] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2014. Engaging with massive online courses. In *Proceedings of the 23rd international conference on World wide web*. ACM, 687–698.
[2] John D Bransford, Ann L Brown, Rodney R Cocking, et al. 2000. *How people learn*. Vol. 11. Washington, DC: National academy press.

Table 4: Within-quartile final exam improvement for each course performance quartile. For each quartile the exam mean and standard deviation are reported for below and above average minutes-viewed.

Course Performance Q_i (Rank)	Threshold τ_i (hours)	Exam $\langle score_i(t < \tau_i)\rangle$	Exam $\langle score_i(t \geq \tau)\rangle$	Mann Whitney p	Peer Improvement Absolute Δ_i	Fractional Reduction $\frac{\Delta_i}{100-\langle score_i\rangle}$
Q_1 (Rank 0% – 25%)	5.6	55.4 ± 16.3	69.9 ± 16.9	0.005	14.4	0.38
Q_2 (Rank 25% – 50%)	6.2	68.6 ± 12.2	78.6 ± 8.6	0.002	10.1	0.39
Q_3 (Rank 50% – 75%)	9.2	74.1 ± 10.4	80.4 ± 7.7	0.008	6.3	0.30
Q_4 (Rank 75% – 100%)	11.3	81.1 ± 10.1	86.2 ± 8.5	0.029	5.2	0.38

[3] Lori Breslow, David E Pritchard, Jennifer DeBoer, Glenda S Stump, Andrew D Ho, and Daniel T Seaton. 2013. Studying learning in the worldwide classroom research into edX's first MOOC. *Research & Practice in Assessment* 8 (2013), 13–25.

[4] Jim Broadbent and WL Poon. 2015. Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education* 27 (2015), 1–13.

[5] Sheryl Burgstahler. 2011. Universal design: Implications for computing education. *ACM Transactions on Computing Education (TOCE)* 11, 3 (2011), 19.

[6] Maria Cardelle-Elawar. 1995. Effects of metacognitive instruction on low achievers in mathematics problems. *Teaching and Teacher Education* 11, 1 (1995), 81–95.

[7] Yuren Xie Biswadeep Chatterjee Hongyu Wang Chirantan Mahipal, Lawrence Angrave and Zhengru Qian. 2019. âĂIJWhat did I just miss?!âĂİ Presenting ClassTranscribe, an Automated Live-captioning and Text-searchable Lecture Video System, and Related Pedagogical Best Practices. In *2019 ASEE Annual Conference & Exposition*. ASEE Conferences, Tampa, Florida. https://peer.asee.org/31926.

[8] Gregory W. Corder and Dale I. Foreman. 2014. *Nonparametric statistics: a step-by-step approach* (second edition ed.). Wiley, Hoboken, New Jersey.

[9] Patricia L Davies, Catherine L Schelly, and Craig L Spooner. 2013. Measuring the effectiveness of Universal Design for Learning intervention in postsecondary education. *Journal of Postsecondary Education and Disability* 26, 3 (2013), 195–220.

[10] Tereza Dean, Anita Lee-Post, and Holly Hapke. 2017. Universal design for learning in teaching large lecture classes. *Journal of Marketing Education* 39, 1 (2017), 5–16.

[11] Heather K Evans and Victoria Cordova. 2015. Lecture videos in online courses: A follow-up. *Journal of Political Science Education* 11, 4 (2015), 472–482.

[12] John H Flavell. 1979. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist* 34, 10 (1979), 906.

[13] Jon Gunderson and Nicholas Hoyt. 2018. AInspector Mozilla Plugin. https://ainspector.github.io/

[14] Paula Jones, Kim Naugle, and M Kolloff. 2008. Teacher presence: Using introductory videos in online and hybrid courses. *Learning Solutions Magazine* (2008), 1–10.

[15] Margaret E King-Sears, Todd M Johnson, Sheri Berkeley, Margaret P Weiss, Erin E Peters-Burton, Anya S Evmenova, Anna Menditto, and Jennifer C Hursh. 2015. An exploratory study of universal design for teaching chemistry to students with and without disabilities. *Learning Disability Quarterly* 38, 2 (2015), 84–96.

[16] René F Kizilcec, Chris Piech, and Emily Schneider. 2013. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge*. ACM, 170–179.

[17] Murali Mani and Quamrul Mazumder. 2013. Incorporating metacognition into learning. In *Proceeding of the 44th ACM technical symposium on Computer science education*. ACM, 53–58.

[18] Majid Moridani. 2007. Asynchronous video streaming vs. synchronous videoconferencing for teaching a pharmacogenetic pharmacotherapy course. *American journal of pharmaceutical education* 71, 1 (2007), 16.

[19] John L Nietfeld, Li Cao, and Jason W Osborne. 2005. Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *The Journal of Experimental Educational* (2005), 7–28.

[20] Sakineh Nikzad, Abbas Azari, Hosseinali Mahgoli, and Nasrin Akhoundi. 2012. Effect of a procedural video CD and study guide on the practical fixed prosthodontic performance of Iranian dental students. *Journal of Dental Education* 76, 3 (2012), 354–359.

[21] Jennifer Parham, Leo Gugerty, and DE Stevenson. 2010. Empirical evidence for the existence and uses of metacognition in computer science problem solving. In *Proceedings of the 41st ACM technical symposium on Computer science education*. ACM, 416–420.

[22] David R Parker, Linda E Robinson, and Robert D Hannafin. 2007. âĂIJBlendingâĂİ technology and effective pedagogy in a core course for preservice teachers. *Journal of Computing in Teacher Education* 24, 2 (2007), 49–54.

[23] S Ramlogan, V Raman, and J Sweet. 2014. A comparison of two forms of teaching instruction: video vs. live lecture for education in clinical periodontology. *European Journal of Dental Education* 18, 1 (2014), 31–38.

[24] Jia Chen Ren, Mark Hasegawa-Johnson, and Lawrence Angrave. 2015. Classtranscribe: a new tool with new educational opportunities for student crowdsourced college lecture transcription. In *ISCA International Workshop on Speech and Language Technology in Education, SLaTE 2015, Leipzig, Germany, September 4-5, 2015*. 179–180. http://www.isca-speech.org/archive/slate_2015/sl15_179.html

[25] Rebecca H Rutherfoord and James K Rutherfoord. 2008. Exploring teaching methods for on-line course delivery-using universal instructional design. In *Proceedings of the 9th ACM SIGITE conference on Information technology education*. ACM, 45–50.

[26] NI Scagnoli and C Packard. 2011. Improving a blended delivery model in a large enrollment business course. *Proceedings of the Academy of Human Resource Development. Chicago, IL: Academic Press* (2011).

[27] Norma I Scagnoli, Jinhee Choo, and Jing Tian. 2019. Students' insights on the use of video lectures in online classes. *British Journal of Educational Technology* 50, 1 (2019), 399–414.

[28] Gregory Schraw. 1998. Promoting general metacognitive awareness. *Instructional science* 26, 1-2 (1998), 113–125.

[29] Daniel T Seaton, Yoav Bergner, Isaac Chuang, Piotr Mitros, and David E Pritchard. 2014. Who does what in a massive open online course? (2014).

[30] Tayfun Tuna, Tuhin Dey, Jaspal Subhlok, and Leigh Leasure. 2017. Video Supported Flipped Classroom. In *Proceedings of EdMedia + Innovate Learning 2017*, Joyce P. Johnston (Ed.). Association for the Advancement of Computing in Education (AACE), Washington, DC, 1159–1171. https://www.learntechlib.org/p/178432

[31] Tayfun Tuna, Jaspal Subhlok, Lecia Barker, Shishir Shah, Olin Johnson, and Christopher Hovey. 2016. Indexed Captioned Searchable Videos: A Learning Companion for STEM Coursework. *Journal of Science Education and Technology* (2016), 1–18. https://doi.org/10.1007/s10956-016-9653-1

[32] Tayfun Tuna, Jaspal Subhlok, Lecia Barker, Varun Varghese, Olin Johnson, and Shishir Shah. 2012. Development and Evaluation of Indexed Captioned Searchable Videos for STEM Coursework. In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education (SIGCSE '12)*. ACM, New York, NY, USA, 129–134. https://doi.org/10.1145/2157136.2157177

[33] Scott W Vanderstoep, Paul R Pintrich, and Angela Fagerlin. 1996. Disciplinary differences in self-regulated learning in college students. *Contemporary educational psychology* 21, 4 (1996), 345–362.

[34] Andria Young and Jane D Fry. 2008. Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning* 8, 2 (2008), 1–10.