Students' Preferences Between Traditional and Video Lectures: Profiles and Study Success

Petri Ihantola
petri.ihantola@helsinki.fi
University of Helsinki
Finland

Juho Leinonen
juho.leinonen@helsinki.fi
University of Helsinki
Finland

Matti Rintala
matti.rintala@tuni.fi
Tampere University
Finland

ABSTRACT

Videos play a key role in modern education. Although the impact and design of video material are relatively well understood, individual differences and patterns, especially in the context of computing education, in video utilization are not. In this work, we study how students attend live lectures and watch video lectures in a data structures and algorithms course. Students had the option to attend the live lectures, watch the lectures as videos, or both. We cluster data related to live lecture attendance and video watching collected over multiple years and find three distinct profiles: 1) active students who prefer lectures, 2) active students who prefer videos and 3) those who provide only limited attention toward online or video lectures. Whereas previous research suggests that watching video lectures is related to at least equal performance when compared to attending live lectures, we found that (in our context) students attending live lectures outperformed all the other groups. This may be related to a selection bias in our setup, however. Thus, the use of video lectures as a part of versatile study material in computing education needs more attention in the future.

CCS CONCEPTS

• **Social and professional topics** → *Computing education.*

KEYWORDS

video lectures, live lectures, traditional lectures, lecture attendance, clustering, performance, retention

1 INTRODUCTION

Videos have become a major part of modern online and blended courses [6]. In online courses, videos can serve as a replacement to traditional live lectures. In blended courses, they can be useful as additional learning content, or in the case of videotaped live

lectures, help students refresh their memory and give students the option to not attend the live lectures but instead watch them later as videos. Moreover, short videos tailored specifically for online education are used in many MOOCs, as they typically result in higher audience retention (and thus better learning) than longer ones [7]. Despite the benefits of short videos, it's still common (especially in non-MOOC settings) to stream traditional lectures as well to let students access study material anytime and anywhere.

The use and research of video materials as part of computing education is active [11]. An ITiCSE 2007 working group, for example, created a repository of videos to support learning of debugging already over ten years ago [14]. The use of practical coding demonstration videos have been reported repeatedly after that (e.g., [1, 15]). These kinds of practical videos are critical especially in illustrating the programming process – a task that is difficult to do in a traditional text book [2]. In addition to technical demonstrations, recordings of more traditional lectures are also used in computing education (e.g., [5, 16, 17]).

Although previous research has explored the effectiveness and design of traditional and video lectures from many angles, the way how students behave if they can freely select whether they attend lectures or watch recorded presentations later is poorly understood. Moreover, the context of many previous studies is different from computing education which relies on interactive or even automatically assessed online content. Such content is more engaging than videos [10], affecting the use of such material and consequently the learning design.

In this study, we observe *what kinds of behavioral profiles are related to lecture participation and watching lectures as videos* (RQ1), and investigate *whether students with different profiles differ with respect to their course performance* (RQ2).

2 RELATED WORK

The use of videos to replace and supplement traditional lectures has been researched for decades. The findings are still in many parts inconclusive. In the following, we will cover the related work in combining video lectures and traditional lectures and then video lectures as a strict alternative to traditional lectures.

2.1 Combining Traditional Lectures and Videos

Wieling and Hofman [18] analyzed the number of times students viewed online video lectures and attended physical live lectures. They found that both live lecture attendance and video lecture watching correlated positively with the exam score, and the effects of both were close to equal. Additionally, they found that for students with low live lecture attendance, watching the video lectures can help them achieve similar results to those with high lecture

attendance. Thus, they suggest that offering videos of live lectures is an easy and efficient way of helping students who are unable to attend the live lectures.

Stephenson [15] recorded code demonstrations done in a CS1 class and provided students the opportunity to watch them later as videos. He found that students who utilized the opportunity found the videos educational, although they also warn of the significant time commitment that comes with making videos. Similarly, Aycock [1] developed a system for recording and playing back live code demonstrations, and found that students' feedback was positive, which is also in line with findings by Bennedsen and Caspersen [2].

Cardall et al. [3] studied a course where students had the option to either attend live lectures or watch video lectures, or both. They found that a majority of students only watched the lectures live and that the main reason for students opting for the live lectures was a lack of motivation to watch the video lectures. However, all students reported having watched at least a single video lecture, and students found the video lectures beneficial for learning.

Vilner et al. [16] studied video utilization in a CS1 course that combined lecture videos and face-to-face meetings. They found that the more videos students watched, the more likely they were to attend the final exam. However, there was no correlation between the number of videos watched and success in the exam; they hypothesize that this is because practice is more important than watching videos (or reading textbooks) for learning programming, which is supported by later studies, for example the ones by Koedinger et al. [10] and Vinaja et al. [17].

Dazo et al. [5] studied how students utilized videos in a flipped CS1 classroom in which students were expected to watch videos before class. They found that based on video usage data, most students did not watch the videos before class and few reviewed content by rewatching videos. However, adding small incentives such as requiring students to post a comment about each video increased students' engagement with the videos. Additionally, when there were no incentives to watch videos, students who nevertheless watched them performed better in the exam.

2.2 Experimental studies between traditional lectures and videos

There have been some studies comparing having only traditional live lectures and having only video lectures [8, 9, 12, 13] with mixed results. Already in 1980, Paegle et al. [12] compared video lectures to traditional lectures. They found that while students' scores to multiple-choice questions given after the lectures were similar between the groups, students rated traditional lectures more highly. Similarly, in a study by Schreiber et al. [13], there were no differences in students' ability as measured by multiple-choice questions given after lectures. Still, students preferred live lectures, although they noted that video lectures can be convenient due to the possibility of stopping, reviewing, and repeating content.

On the other hand, some studies have found students preferring video lectures. Halupa and Caldwell [8] compared two different student cohorts, where one cohort only had live lectures and the other were given supplemental videos in addition to the live lectures. They found that there were no differences between the groups with regards to exam scores, but that the group with

videos found the videos helpful. In a study by Kestin et al. [9] in an introductory physics course, students were randomly split into two groups, where one group had live demonstrations and the other video demonstrations. They found that students who watched the video demonstrations performed better in multiple-choice questions related to the demonstration, and that students' enjoyment of the demonstrations was equal.

Based on the related work presented here, having both live lectures and videos of those lectures seems to be the best option as most studies found that video lectures are at least as beneficial for learning as live lectures [8, 9, 12, 13, 18], even though students might prefer live lectures [3, 12, 13].

3 METHODS

3.1 Context

The participants of this study were 434 students taking the data structures and algorithms course taught at Tampere University of Technology between 2016 and 2018. This 5 ECTS[1] course is targeted primarily for the second year Bachelor students and offered once a year.

The following explains how the course was organized and graded. Throughout the course students were able to collect points to three "buckets":

Theory Points from lecture essays submitted before weekly lectures (10p/essay), being present in lectures (20p/lecture), or watching lectures on the web (10p/lecture) were added together. Getting points from a single lecture by both being present and watching the video was not possible. Students had to watch at least a third of the web lecture within four weeks from the lecture date to get the points. The weight of this bucket in final grading was 25%.
Constant practising Points from weekly manually graded small tasks (10p/question), being present in weekly exercise sessions (10p/session), and points from automatically graded algorithm simulation exercises (implemented by using the JSAV library) were added together. The weight of this bucket in final grading was 35%.
Applying knowledge Three larger programming assignments, each having an equal weight contributed 40% in the final grade.

In addition to these activities, there was an exam at the end of the course that was graded as passed or failed. Students had to pass this exam in order to pass the course. Although the final grading of the exam was coarse grained, students were able to get fine-grained feedback from the exam and all the tasks within the exam were graded individually.

3.2 Measures

For each student, we had the fine-grained grading data described in the previous section. In this study, we used the following subset of that data. The percentage of lectures a student participated (i.e., Lecture count) and the percentage of lectures they watched from videos (i.e., Video count). Students' performance was triangulated by using the following measures: exam retention (i.e., whether students participated in the exam or not), exam points (i.e., fine-grained points from the exam), and the grade of the "applying knowledge"

---

[1]European Credit Transfer and Accumulation System

Table 1: Mean values, standard deviations and one-way ANOVA of original measures between all course instances.

	2016		2017		2018				
	M	SD	M	SD	M	SD	F	p	η^2
Lecture count	0.515	0.349	0.526	0.351	0.437	0.377	$F(2,475) = 2.922$	.055	.012
Video count	0.292	0.265	0.235	0.246	0.204	0.238	$F(2,475) = 4.856$	.008	.008
Exam	12.917	2.867	11.992	2.747	12.055	2.669	$F(2,430) = 5.04$	.007	.023
Programming	2.930	0.928	2.970	1.141	2.987	1.139	$F(2,441) = 0.101$	.904	.000

category. The measures were selected so that they illustrate retention and the mastery of the course topics both from a theoretical and a practical point of view. Measures that were (partially) derived based on attendance were excluded. Moreover, algorithm simulation exercises were excluded as although arrangements between the course versions were almost identical, there were small differences at which point of the course the simulation exercises were released and when their deadlines were. Yearly statistics together with one-way ANOVA to estimate between-years differences in the measures are provided in Table 1.

For further analysis, data from different years were combined. Descriptive statistics and correlations between all the variables in the combined data set are provided in Table 2.

As illustrated in Table 1, there were some small, but significant differences (between years) in some of the measures. Thus, two versions of the combined data set were used. One where the raw values were combined and another where variables were normalized for each year and the normalized z-scores were then combined.

Table 2: Pearson correlations between the continuous variables in the combined data set with means (M) and standard deviations (SD). * p<.05, **p<.01, and ***p<.001

|	1.	2.	3.	M	SD
1. Lecture count				0.529	0.351
2. Video count	-0.57***			0.257	0.256
3. Exam	0.28***	-0.07		12.31	2.79
4. Programming	0.32***	0.06	0.41***	2.96	1.07

3.3 Analysis

To identify the potential profiles in the lecture attendance and video watching (RQ1), we clustered the data based on lecture and video counts by using the hierarchical clustering approach. The number of clusters was based on the the NbClust [4] R-package, which combines 30 different indices and selects the number of clusters based on the majority vote. Chi-square test of independence was used to see if course versions are evenly distributed between the clusters (we refer to the clusters as "profiles").

Two-way analyses of variance (ANOVA) were performed to identify whether the profiles and years differ in terms of their performance (RQ2). The interaction of years and profiles was also examined. A Tukey's Honestly Significant Difference test was used as the related post-hoc test. All of the analyses were performed separately for the raw data and normalized data as described in the previous section.

4 RESULTS

4.1 Profiles

Raw-data and per-year normalized data both indicated a three cluster solution (see Figure 1). Scatter plots illustrating the resulted profiles are provided in Figure 2. The median silhouette widths of the solution were 0.51 and 0.47, respectively for the raw data and normalized approaches, indicating a reasonable fit of the models. In the future, we will refer to these cluster as high-lecture, high-video, and low-all. Although the shapes of the clusters are similar, there are differences in where the lines are drawn. Sizes of the clusters are (101,260,117) (for raw data) and (163,251,64) (for normalized) for low-all, high-lecture, and high-video respectively.
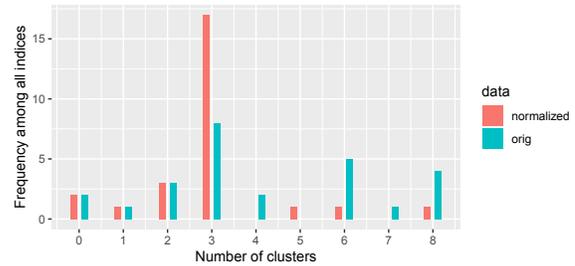
Figure 1: Frequency among all indices used to decide the number of clusters.

In the raw data based approach, A chi-square test of independence showed that there was a significant association between the course versions and profiles ($\chi^2(4) = 19.389$, $p = .001$). The significant differences are that the 2016 and 2017 course versions were underrepresented in the first cluster with standardized residuals of -2.3 and -2.1, respectively for 2016 and 2017. Correspondingly the 2018 course version was over-represented in the first cluster (std.residuals=4.4) and underrepresented in the second cluster (std.residuals=-2.6). There was no such association in the normalized approach ($\chi^2(4) = 1.4236$, $p = .840$).

4.2 Performance

Table 3 summarizes performance statistics for the clustering based on the unchanged data. The statistics related to the normalized approach are similar. Almost all of the students who were in either the high-lecture or the high-video clusters attended the exam, but around a third of the students in the low-all cluster did not attend the exam. The differences are statistically significant (($\chi^2(2) = 111.59$, $p < .000$) so that the low-all group is clearly underrepresented in exam takers (std.residuals=-10.5). Correspondingly the
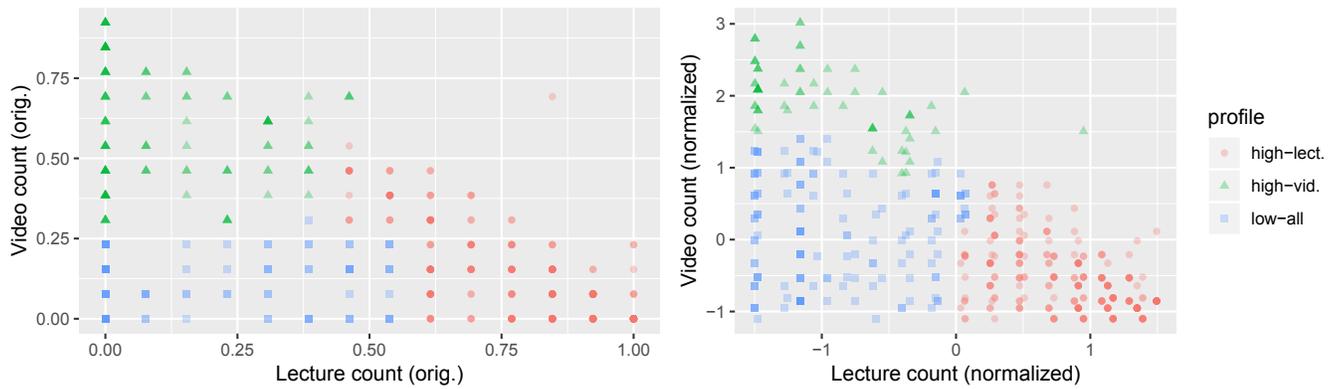
Figure 2: Profiles related to lecture attendance and video watching. Overlapping values are illustrated with darker color.

high-lecture group is clearly over-represented (std.residuals=6.4) and the high-video group is slightly over-represented (std.residuals=2.4) among the exam takers.

Table 3: Study success between profiles.

	Exam			Programming		
profile	n	M	SD	n	M	SD
high-lecture	256	12.95	2.40	260	3.26	0.97
high-video	113	11.47	2.97	117	2.75	1.10
low-all	64	11.24	3.22	101	2.23	0.98

There was a significant effect of profile on the exam with both clustering approaches. Moreover, the effect of profile on programming grade was significant with the original data but not significant with the normalized approach. There were no significant effects on course version or interaction between course version and profile. Details of the test statistics are provided in Table 4.

Post-hoc comparisons using Tukey's HSD test indicated that the exam points for the low-all condition were significantly lower than for the high-lecture condition with both clustering approaches. Effect sizes (Cohen's d) were 0.60 and 0.62, respectively for the original and normalized approaches. Clustering based on the original data lead to situation where high-video had lower exam scores than high-lecture, and there were no significant differences between low-all and high-video. In the normalized approach, high-video performed better than low-all, but the differences between high-lecture were no longer significant. The effect sizes where the high-lecture group outperformed high-video were 0.54 and 0.49, respectively for exam and programming scores with the original data.

All the pairwise differences are provided in Table 5 and illustrated in Figure 3. In the case of the normalized data, the statistically non-significant differences related to programming scores are omitted.

5 DISCUSSION AND CONCLUSIONS

We have analyzed log data and study records of over 400 students' data from three years. The results indicate that based on lecture participation and logs from the video platform, students can be clustered into three groups: 1) active students who prefer lectures

Table 4: Two-way Anova with cluster and year as independent variables to explain differences in exam and programming perf. separately for both clustering solutions.

		Sum Sq Type III	Df	F	Pr(>F)
Exam (original)	year	27.75	2	1.95	0.1437
	profile	66.93	2	4.70	**0.0096**
	year:profile	11.09	4	0.39	0.8163
	Residuals	3018.81	424		
Programming (original)	year	0.70	2	0.34	0.7107
	profile	12.16	2	5.97	**0.0028**
	year:profile	1.64	4	0.40	0.8066
	Residuals	443.23	435		
Exam (normalized)	year	0.33	2	0.17	0.8439
	profile	4.99	1	5.14	**0.0238**
	year:profile	0.28	2	0.14	0.8676
	Residuals	413.85	427		
Programming (normalized)	year	1.94	2	1.02	0.3631
	profile	3.13	1	3.27	0.0712
	year:profile	1.89	2	0.99	0.3727
	Residuals	419.18	438		

(high-lecture), 2) active students who prefer videos (high-video) and 3) those who provide only limited attention toward online or video lectures (low-all).

Students in the low-all group were more likely to drop out of the course. Moreover, among those who remained until the end of the course, students in the high-lecture group performed slightly better in the final exam and programming tasks throughout the course. Differences in the exam performance seem to be related to retention as well. Students who prefer traditional lectures are likely to attend nearly all of the lectures while students preferring video lectures are likely to skip nearly a fourth of the lecture teaching, even if they continue with other activities till the end of the course.

Our study comes with a set of limitations, however. Most importantly, although both lecture and video participation are rewarded, traditional lectures contribute more to the final grade. Thus, there
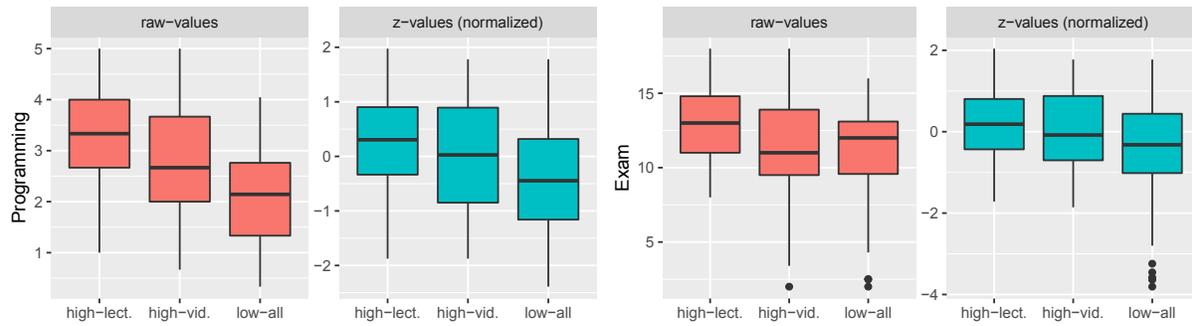
Figure 3: Programming assignment and exam performance between clusters.

Table 5: Tukey's HSD post-hoc test for ANOVA in Table 4.

	compared clusters	diff	conf. low	conf. high	p (adj.)
Exam (original)	high lect - low all	1.71	0.82	2.59	0.00
	high vid - low all	0.23	-0.76	1.22	0.85
	high vid - high lect	-1.47	-2.19	-0.76	0.00
Prog. (original)	high lect - low all	1.02	0.71	1.34	0.00
	high vid - low all	0.51	0.16	0.87	0.00
	high vid - high lect	-0.51	-0.77	-0.24	0.00
Exam (normalized)	high lect - low all	0.63	0.38	0.88	0.00
	high vid - low all	0.44	0.09	0.79	0.01
	high vid - high lect	-0.19	-0.51	0.13	0.33

might be a selection bias so that students preferring video lectures are inherently less motivated. Internal validity of our clustering is also worth looking at. Although courses were designed to be similar, there was a danger that inevitable small differences between the course versions affect the clustering. Therefore, we examined also an alternative clustering approach where all the measures were scaled within the course version from where they were collected. Although the internal validity of this second model raises some concerns (silhouette score < 0.5), the finding from both approaches support each other, increasing our trust that the profiles are suitable. Moreover, the role of the course versions was also taken into account when analyzing programming grades and exam points, and it did not explain the differences between the clusters.

Our findings can be interpreted in two ways. First, the role of traditional lectures is still important and transforming this online may have even negative consequences; When students are given the opportunity to view lectures online it's easier to slack off or even drop out of the course. On the other hand, it may be that students who preferred videos would have not been able to participate in the course at all if videos were not available. More research is needed to better understand who watch video lectures and for whom video lectures should be targeted when traditional courses are taught (partially) online. Moreover, more research is needed to understand the role of videos in complex learning environments (such as computing) where multiple online elements with varying levels of interaction compete for students' attention.

REFERENCES

[1] John Aycock. 2018. Stick to the script: Lightweight recording and playback of live coding. In *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education.* 350–351.

[2] Jens Bennedsen and Michael E Caspersen. 2005. Revealing the programming process. In *Proceedings of the 36th SIGCSE technical symposium on Computer science education.* 186–190.

[3] Scott Cardall, Edward Krupat, and Michael Ulrich. 2008. Live lecture versus video-recorded lecture: are students voting with their feet? *Academic Medicine* 83, 12 (2008), 1174–1178.

[4] Malika Charrad, Nadia Ghazzali, Véronique Boiteau, Azam Niknafs, and Maintainer Malika Charrad. 2014. Package 'nbclust'. *Journal of statistical software* 61 (2014), 1–36.

[5] Suzanne L Dazo, Nicholas R Stepanek, Robert Fulkerson, and Brian Dorn. 2016. An empirical analysis of video viewing behaviors in flipped CS1 courses. *ACM Inroads* 7, 4 (2016), 99–105.

[6] Sue Greener. 2015. Flipped or Blended? What's the Difference and Does it Make a Difference to Learning in HE?. In *International Conference on e-Learning.* Academic Conferences International Limited, 146.

[7] Philip J Guo, Juho Kim, and Rob Rubin. 2014. How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning@ scale conference.* 41–50.

[8] Colleen M Halupa and Benjamin W Caldwell. 2015. A Comparison of a Traditional Lecture-Based and Online Supplemental Video and Lecture-Based Approach in an Engineering Statics Class. *International Journal of Higher Education* 4, 1 (2015), 232–240.

[9] Greg Kestin, Kelly Miller, Logan S McCarty, Kristina Callaghan, and Louis Deslauriers. 2020. Comparing the effectiveness of online versus live lecture demonstrations. *Physical Review Physics Education Research* 16, 1 (2020), 013101.

[10] Kenneth R Koedinger, Jihee Kim, Julianna Zhuxin Jia, Elizabeth A McLaughlin, and Norman L Bier. 2015. Learning is not a spectator sport: Doing is better than watching for learning from a MOOC. In *Proceedings of the second (2015) ACM conference on learning@ scale.* 111–120.

[11] Andrew Luxton-Reilly, Ibrahim Albluwi, Brett A Becker, Michail Giannakos, Amruth N Kumar, Linda Ott, James Paterson, Michael James Scott, Judy Sheard, and Claudia Szabo. 2018. Introductory programming: a systematic literature review. In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education.* 55–106.

[12] RD Paegle, EJ Wilkinson, and MB Donnelly. 1980. Videotaped vs traditional lectures for medical students. *Medical education* 14, 6 (1980), 387–393.

[13] Benjamin E Schreiber, Junaid Fukuta, and Fabiana Gordon. 2010. Live lecture versus video podcast in undergraduate medical education: A randomised controlled trial. *BMC medical education* 10, 1 (2010), 1–6.

[14] Beth Simon, Sue Fitzgerald, Renée McCauley, Susan Haller, John Hamer, Brian Hanks, Michael T Helmick, Jan Erik Moström, Judy Sheard, and Lynda Thomas. 2007. Debugging assistance for novices: a video repository. *ACM SIGCSE Bulletin* 39, 4 (2007), 137–151.

[15] Ben Stephenson. 2019. Coding Demonstration Videos for CS1. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education.* 105–111.

[16] Tamar Vilner, Ela Zur, and Ronit Sagi. 2012. Integrating video components in CS1. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education.* 123–128.

[17] Robert Vinaja. 2014. The use of lecture videos, ebooks, and clickers in computer courses. *Journal of Computing Sciences in Colleges* 30, 2 (2014), 23–32.

[18] MB Wieling and WHA Hofman. 2010. The impact of online video lecture recordings and automated feedback on student performance. *Computers & Education* 54, 4 (2010), 992–998.