

Objectives

- Investigate how use of online quiz tools *during* asynchronous video lectures (henceforth AVL) affects learning. I study both...
 - ...the effect on actual learning compared to receiving a long quiz in the end of the AVL
 - ...how quizzing (during or after AVL) influences students' perceived learning
- I do this by conducting a randomized control trial (henceforth RCT) among 1st-year students in the Computer Science-Economics (CompEcon) program at UCPEH.

Introduction

During more than a year teaching at universities has moved online due to the COVID-19 pandemic. AVL has removed an essential element of lectures, namely students' chances to get instant feedback on their level of understanding by asking questions.

The question is whether it is possible to replace the oral instant feedback with something else. One such something is online quizzes. [3] document that use of tools in lectures to activate students increases student learning, which also [7, 4, 2] underline. [6, 1, 5] focus specifically on the online or hybrid classroom and report positive effects on student learning and motivation from using well-designed online tools like quizzes. However, none of these studies investigate where to optimally place the quiz during the online lecture.

The aim of this project is to provide guidance to lecturers on whether and, if so, how use of quizzes can support learning in the asynchronous classroom. Even in post-pandemic times it is highly relevant to consider how teaching outside the lecture hall can be improved so that confrontation hours between lecturer and students are spent efficiently.

Randomized Control Trial

I conducted an RCT among 69 students enrolled in the 7 weeks' course *Introductory Probability Theory and Statistics* in the CompEcon program. 4/6 weekly lecture hours were pre-recorded videos and the remaining 2 hours live on Zoom. To be eligible for the exam, students must pass 3/4 home assignments. The RCT took place during lecture 11 which consisted of 7 videos of 2-25 minutes each (average of 12 minutes). Each student was randomly assigned to one of three groups:

- Group 1 17 quiz questions spread out during the videos ($N = 22$)
- Group 2 17 quiz questions after the videos ($N = 24$)
- Group 3 No quiz ($N = 23$)

Before being assigned to groups, I asked the students if they usually watched the AVLs together with others. 2 students reported they did and were consequently (non-randomly) assigned to the same random group. All students then received a group-specific email in their KU inbox with a link to the play list of the AVL of their group and instructions for the lecture. It was explained that during the play list students would be asked to answer questions accessed via a QR-code and that it was *mandatory* to answer all questions within 4 days to qualify for passing home assignment 3 (which was otherwise unrelated to the quiz). This was done to minimize the risk of selection on the participation dimension. 7 students did not meet the criteria, but had all been inactive in the course so far and hence presumably dropped out. There were no requirements to obtain a minimum number of points (each correct answer gave 1 point) as I wanted to mimic the usual teaching situation, not an exam situation.

Group 3 would only see a QR-code to a survey of their experience with the lecture. So would groups 1 and 2 in addition to questions related to the material covered in the AVL. I did not reveal where in the videos the QR-codes would be, as I wanted to motivate all groups to actually watch the AVL. For groups 1 and 2, the email also explained that in-process feedback has potentially positive learning benefits and for all groups the purpose of the intervention was mentioned. Finally, the importance of using their personal link to the AVL was emphasized.

Learning Results

#1 Group 2 performed better: Figure 1 displays the raw average points received by groups 1 and 2 in the quiz. It also informs that across all groups, students watched 5.5 of the 7 videos on average, most likely indicating that the majority of students watched the first 6 videos covering the curriculum (video 7 was titled as a recap and from logfiles of views of previous AVLs on Absalon, students tended to skip the short recap videos). According to the figure, group 2 performed better, and Table 1 regresses the total points on group membership in addition to other controls. Further models have been tested, e.g. with interactions between all or some variables, but these additional controls showed up as insignificant and are therefore not reported. The group 2 variable is significantly positive in all specifications at least at the 10% level. This indicates that students in group 2 obtained 1.7-2.1 more points on average than group 1 students conditional on the other controls. Since group assignment was perfectly random and there is no evidence of attrition bias, this can be interpreted as a causal effect. The effect corresponds to achieving 10-12.4% more correct answers, something that, if the effect is scalable to a final exam format, could imply a 1 grade difference in the course grade on the 7-point grading scale. Whether this is a lasting and scalable effect and will influence exam results cannot be concluded from this experiment though.

#2 Watching alone had positive effect for group 1: The only other variable that has a significant effect is the share of the AVL which the student watched alone, cf. column (3) of Table 1. Column (4) allows for a heterogeneous effect of the share by group and clarifies that the positive effect of watching more of the AVL alone only exists for group 1, whereas it has a negative effect for students in group 2. A possible explanation is that students in group 1 who watch the AVL in groups may feel a pressure to answer quiz questions fast so they do not delay their peers as the video should not be continued until the question was answered. For group 2 students, on the other hand, they may benefit from discussing the quiz questions in the end with peers, but may not experience the same pressure to finish fast since everyone is filling out the quiz *after* the AVL. Despite this negative effect for group 2, the combined effect of being assigned to group 2 instead of group 1 is still positive, ceteris paribus. 90% of videos were watched alone in all groups though, cf. light blue bar in Figure 1, and 90% of students watched all videos alone (not shown), so results may be sensitive to mis-reporting of the share.

#3 Time use does not explain group 2 effect: Even though group 1 students spent more time on the quiz in total or per quiz question according to Figure 1, they did not receive more points. Minutes spent is a questionable control since spending more time can either be because they are more motivated to do well on the quiz (positive effect) or they may find the questions harder to answer than group 2 members do (negative effect). The sign of the effect is therefore not clear ex ante. This reduction-form analysis cannot disentangle the two contributions. It can only provide an estimate of the combined effect which is statistically and economically insignificant, cf. columns (6) and (7) of the table. This is an indication that the two suggested channels may cancel each other out.

#4 Preparation does not explain group 2 effect: For preparation there is no statistically significant effect either, cf. column (5). Again this variable may be endogenous since students who prepare everything may do so because they are very motivated (positive effect of 'AVL') or because they find the subject really difficult (negative effect of 'AVL'). Interactions between the preparation variable and group 2 variable is insignificant (not shown). The blue bars in Figure 2 indicate that about 50% of students in all three groups prepared 'some', though this share is slightly higher for group 1, which also has a higher tendency to prepare everything. The only statistically significant difference between groups and the decision on how much to prepare in advance, is on the probability of choosing 'all' relative to the probability of choosing 'some' preparation. This relative risk ratio is negatively affected if being assigned to group 2 instead of group 1 (results from multinomial logit, not shown). This should not be a threat to identification of the group 2 effect though, since the email instructions for group 1 and 2 were identical except for the link. Hence, it is unlikely that group 2 members have a lower tendency to prepare everything due to an endogenous response to their group assignment. In any case, the preparation control does not affect the conclusion that the

group 2 effect is statistically and economically significant.

Table 1: OLS of total quiz points for group 1 and 2

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|-------------------------------|---------|---------|---------|---------|---------|---------|---------|
| Group 2 (base: Group 1) | 1.72*** | 1.52** | 1.88*** | 1.61** | 2.10** | 2.02** | 1.84** |
| # videos watched | (0.426) | (0.425) | (0.485) | (0.411) | (0.657) | (0.657) | (0.615) |
| Share watched alone | | | 0.23** | | | | |
| Group 2 × Share watched alone | | | 0.23** | 0.16** | | | |
| Preparation (base: None) | | | | | | | |
| None | | | | | 4.33 | | |
| All | | | | | 2.13 | | |
| # min. prep per quiz | | | | | | 0.07 | |
| # min. prep on quiz | | | | | | | 0.18 |
| Constant | 7.99*** | 6.98*** | 4.72*** | 7.00*** | 7.97*** | 7.97*** | 7.97*** |
| R ² | 0.075 | 0.123 | 0.130 | 0.129 | 0.129 | 0.130 | 0.131 |
| N | 60 | 60 | 60 | 60 | 60 | 60 | 60 |

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

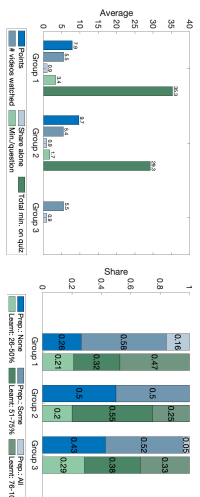


Figure 1: Share and preparation

#5 Group 1 highest perceived learning, group 3 least: As the green bars in Figure 2 reveal, group 1 students had a higher tendency to report a 70-100% understanding of the AVL than any other group (other options are 0-25%, 26-50%, 51-75%). As many as 29% of group 3 only reported 26-50% understanding, a share approx. 10 pp. higher than for the others. However, as Figure 3 shows, the difference in shares across the responses are not statistically significant (from 1) when estimated in a multinomial logit model. 70-100% understanding is the baseline response, i.e. relative risk ratios (RRR) displayed in the figure are all relative to the probability of reporting 70-100%. Additionally controlling for differences in preparation (Figure 4) or number of videos watched (Figure 5) does not affect the estimates of the group effects. The insignificant means the *conclusions on perceived learning cannot be generalized to the population level*, but should only be interpreted with respect to this particular sample.

#6 Group 3 may not realize their lack of understanding: With that in mind, if a group 1 student were to be assigned to group 2 instead, the RRR of understanding 26-50% would be 1.8 times higher, while the RRR of 51-75% would be 3.3 times higher, i.e. he would be more likely to not understand 70-100%. Were he rather assigned to group 3, the RRR would increase by 1.9 and 1.7, respectively. I.e. being assigned to group 3 pushes students away from the highest understanding to a lower degree than if assigned to group 2. This possibly counter-intuitive result may be explained by group 3 students not getting challenged to test their understanding in a quiz and thus are over-confident, whereas group 2 realizes there are parts they did not fully get.

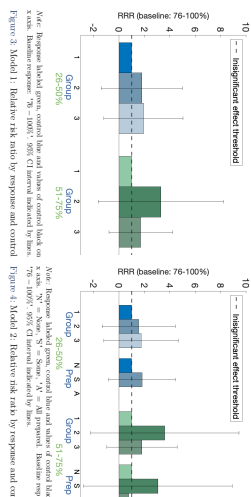


Figure 2: Perceived learning and preparation

Survey Results

After watching the AVL, all groups were asked to fill out a survey on their opinion on the lecture and use of quizzes. Figure 6 shows the distribution of answers by group for the following five questions or statements

- Q1 Quizzes are helpful for my understanding of the AVL
- Q2 I prefer no quizzes so the AVL is shorter
- Q3 Quizzes improve my chances to do well at the exam
- Q4 I prefer quizzes during the AVL rather than after (only group 1 and 2)
- Q5 Quizzes during the AVL make my concentration [...]

Overall, the **attitude towards use of quizzes in AVL is positive across all groups** as the majority (dis)agree or strongly (dis)agree on Q1 (Q2) though a significant share is also neutral. Likewise, most students believe quizzes will improve their performance at the exam (Q3).

On the placement of the quiz either during or after the AVL, there is no clear answer. 42% of group 1 prefers quizzes during the AVL, while this number is only 25% for group 2. This could indicate a bias towards its own recent (mainly positive) experience with the relevant quiz format. The take-away thus is that quizzes in any format is positively perceived by the students, 36% and 45% of group 1 and 2, respectively, were neutral about the format after all.

Especially group 1 found that the quiz during the lecture improved their concentration (68%) while only 5% found it distracting. Group 2 was more indifferent and group 3 more divided on this question as 24% thought it would have worsened their concentration, whereas 43% believed it would have improved it. However, it is important to notice that those who *actually* tried out the in-process quiz format mainly found it beneficial, so their responses should probably weigh more. A few of the videos in the AVL did not contain quizzes for group 1 either, so they should be capable of evaluating the benefit of quizzes during a specific video.

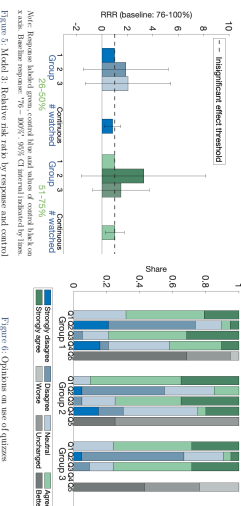


Figure 3: Opinions on use of quizzes

Recommendations

Lecturers are recommended to implement quizzes in their asynchronous lectures as students have a very positive approach to them. Based on the result that actual learning is higher for students who received the quiz in continuation of the AVL and the no clear preference for quizzes during, the main part of questions should be placed after the videos. Importantly though, as students who actually had questions during the AVL found them helpful for their concentration, it is worth adding elements (not necessarily a quiz) that invite them to pause and reflect during the AVL to break the passive 'TV watching'.

References

[1] P. Calders and D. S. Dawson, "The Effect of This Special Online on Student Achievement in Online Economics and Finance Courses," *The Journal of Economic Education* 43(2011). [2] C. Dobbins and P. Duran, "An Evidence-Based Investigation into the Use of Mobile Technology in Enhancing Student Engagement," *In TechTrends* 6(1) (2017). [3] S. Freeman et al., "Active learning increases student performance in science, engineering, and mathematics," *In Proceedings of the National Academy of Sciences* 111(2014). [4] R. H. Key and A. Leberg, "A strategic assessment of audience response systems used in higher education," *In: Association Journal of Educational Technology* 25(2009). [5] A. Bross et al., "Learning and instruction in the hybrid virtual classroom: An investigation of students' engagement and the effect of quizzes," *In Computers and Education* 143(2020). [6] F. Schiller et al., "Comparative Study of Learning Outcomes for Online Learning Platforms," *In Working paper* (Apr. 2021), arXiv:2104.07693 [Cs, Cv]. [7] C. S. Walker et al., "Students taught by live-streamed instruction using active learning teaching strategies outperform students taught by a fully-regarded traditional instructor," *In Working paper* (Apr. 2020), arXiv:2004.09854 [Physics, arXiv:2004.09854].