The use of digital components in an undergraduate statistics course

Florian Ertz and Ralf Münnich
Trier University, Germany
ertz@uni-trier.de

In a research project, the Economic and Social Statistics Department of Trier University was able to get an insight into the use of screencasts and e-tutorials in an undergraduate statistics course. The empirical data show that students make ample use of the flexibility those tools offer. We also see that active exercises are very important in terms of exam preparation. The participation rates and the tenacity in working exercise problems can be improved. A natural next step would be the implementation of adaptive e-learning scenarios.

DIGITAL COMPONENTS DEPLOYED, DLISA, AND COMPOSITION OF THE GROUPS

The Economic and Social Statistics Department of Trier University offers an undergraduate statistics course that introduces students of mainly economic sciences, social sciences, and business informatics to the basic concepts of descriptive and inferential statistics, as well as to the foundations of the statistical programming language *R*. As a pioneer at Trier University, the department enhanced the course with electronic tutorials (e-tutorials) and electronic exams (e-exams) in the winter term 2010/11 in the context of the project *eLearning Infrastructure and Teaching Environment (eLITE)* (Münnich et al., 2011; Höfler-Hoang et al., 2020). Both digital components use the open source solution *ILIAS*. Several e-tutorials, i.e. collections of exercise problems, are rolled out throughout the lecture period and are then available online for students 24/7 until right before the e-exam takes place after the lecture period. Students can practise the application of the statistical methods taught in the lecture by filling in numerical gaps, using drop-downs to respond to single-choice questions, mapping fitting concepts, etc. As the e-tutorials are randomised, i.e. use new subproblems and/or data sets/numbers in successive passes, they offer a deeper opportunity for practice than traditional classroom tutorials. The e-exam can be thought of as a large e-tutorial that is limited to a single pass at a specified time in a computer lab on the premises of the university. The COVID-19 pandemic completely forced university teaching to the digital realm in the spring of 2020. As a replacement for classroom lectures, the department produced screencasts as learning videos. In these, the lecturer or tutor switches between slides, *RStudio*, and handwritten notes on a tablet. Screencasts for the lecture as well as for the traditional tutorial were first rolled out in the summer term 2020, also being available online 24/7. In the summer term 2021, the department initiated the still on-going research project *Digitale Lehr-/Lern-Instrumente in der Statistik-Ausbildung (DLISA)* to investigate how students of the statistics course actually engage with the digital components. For those students who actively consented to the use of their data, we were able to analyse basic information gathered in the various teaching/learning software infrastructures. The collected (student-level) information related to the digital components includes the number of streams per video, total streaming time per video, date and time of first and last access to a video, achieved completion rate per subproblem per e-tutorial and attempt, total time working on an e-tutorial, date and time of first and last attempt. We present results for the first wave of DLISA, i.e. the summer term 2021, here. Up to today, the General Data Protection Regulation does not allow us the data collection/use without an explicit prior consent. As a self-selection of some sort is expected in our setting, we have to acknowledge that our results will not necessarily be representative for the whole group of students taking the statistics course. Using the basic course participant information from our main learning management platform as our *census* (N = 640), we can get an impression of the composition of our sample (n = 69; approx. 10.8 percent). The sample under investigation is smaller when we incorporate exam results later on (n = 27), as not all participants took the e-exam and a small number of participants only took partial exams for descriptive or inferential statistics, respectively. As Figure 1 shows, female students are strongly overrepresented (approx. 59 percent in the sample vs. 45 percent in the overall student group). The distribution across main field of study groups shows a slight over- and underrepresentation of economic science and social science students, respectively. While there are stronger discrepancies for computer science students and students of other fields of study, those two categories only make up a small share of the course participants.

In: Birk, L., Loth, G., Jotzo, L., Binder, K., & Frischemeier, D. (Eds.) (2026). *Statistics and Data Science Education in STEAM. Proceedings of the Satellite Conference of the International Association for Statistical Education (IASE), September/October 2025, Münster, Germany*.

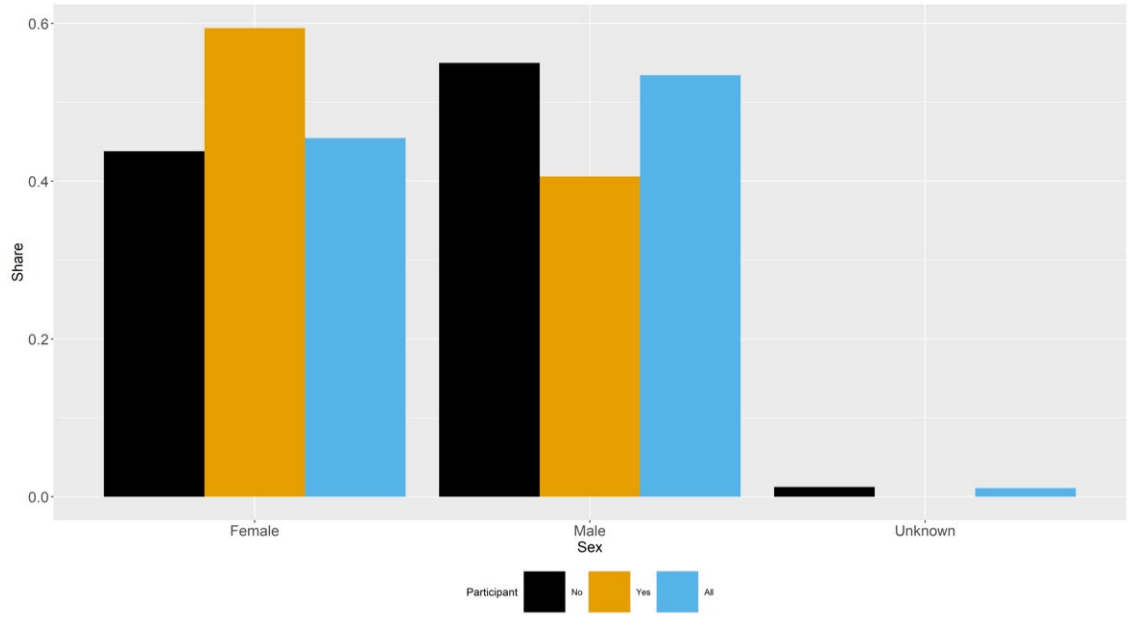

Figure 1. Group comparison – Sex.

ANALYSIS OF THE SCREENCAST DATA

The self-selection is also seen when looking at the development of the participation rates for the screencasts for the lecture and traditional tutorial, respectively, along the lecture period in Figure 2. The participation rates for the whole group can be computed using the accessible number of unique viewers of the individual videos. While we see the familiar decline in participation, the project group (i.e. our sample) keeps engaging with the contents longer. It is important to note that the summer term 2021 was still completely taking place *online* and the pronounced decline might also reflect a certain *pandemic fatigue*. Participation rates in the earlier chapters of the course are lower for the tutorial screencasts than for the lecture screencasts. As Figure 3 indicates, if students start to stream lecture screencasts (i.e. individual videos) of a chapter, the majority of them also finishes the complete chapter materials. The same does not go for the tutorial screencasts, where the median of the completion rate is as low as 50 percent for one of the chapters.

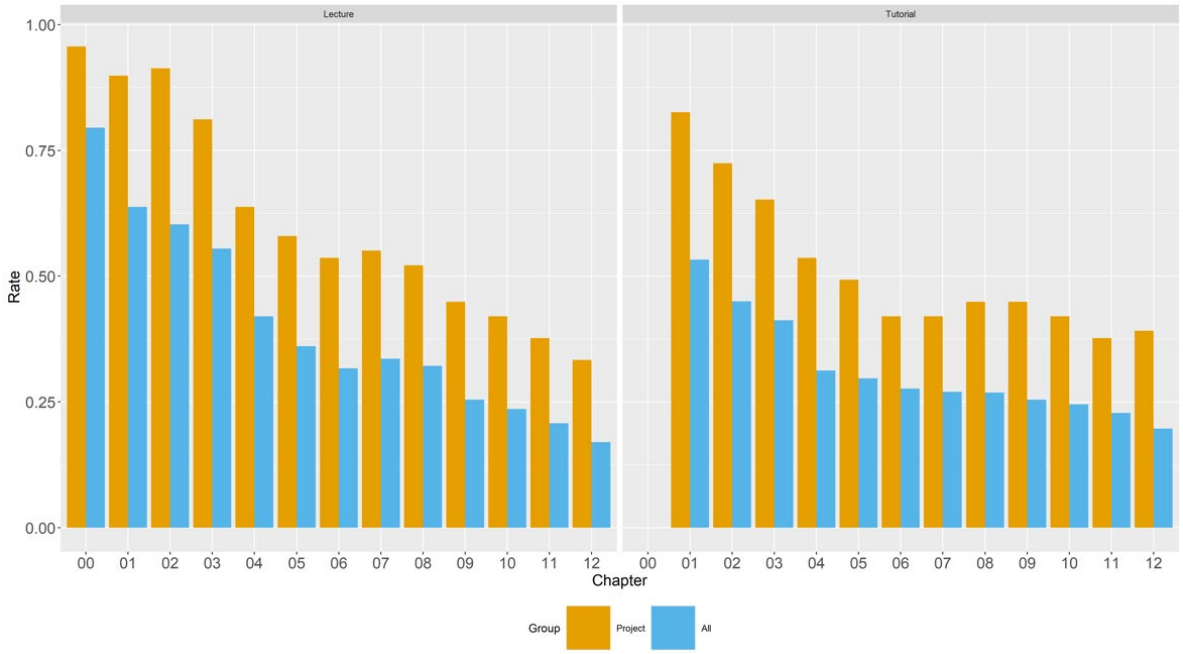

Figure 2. Screencasts – Participation rates.

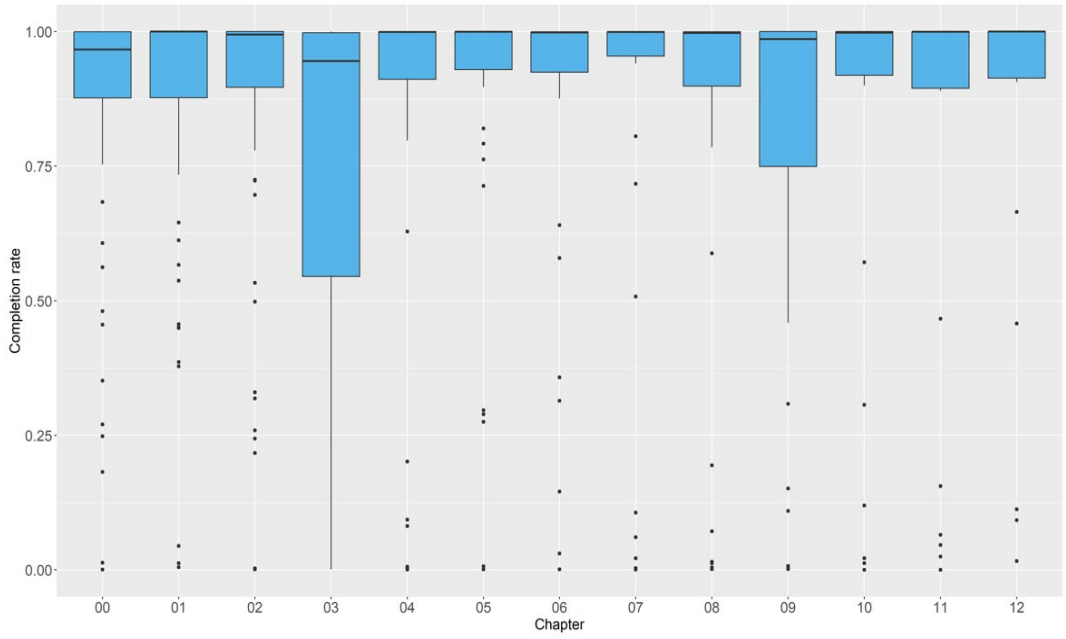

Figure 3. Lecture screencasts – Completion rates.

The time series of the sum of streamed minutes of lecture screencasts per day (Figure 4) shows marked spikes right at the beginning of the lecture period, where *R* is introduced, and around the time that the screencasts introducing estimation and statistical hypothesis testing are rolled out (15 June 2021). This might, not surprisingly, indicate that the lecture slides are not enough in these comparatively demanding parts of the lecture and students need additional explanations. The same time series for the tutorial screencasts shows a different pattern and two pronounced spikes right before the e-exam on 10 August 2021. This pattern is extreme in the corresponding time series for the e-tutorials. This may indicate a shift towards exercises during the final *sprint* before the e-exam.

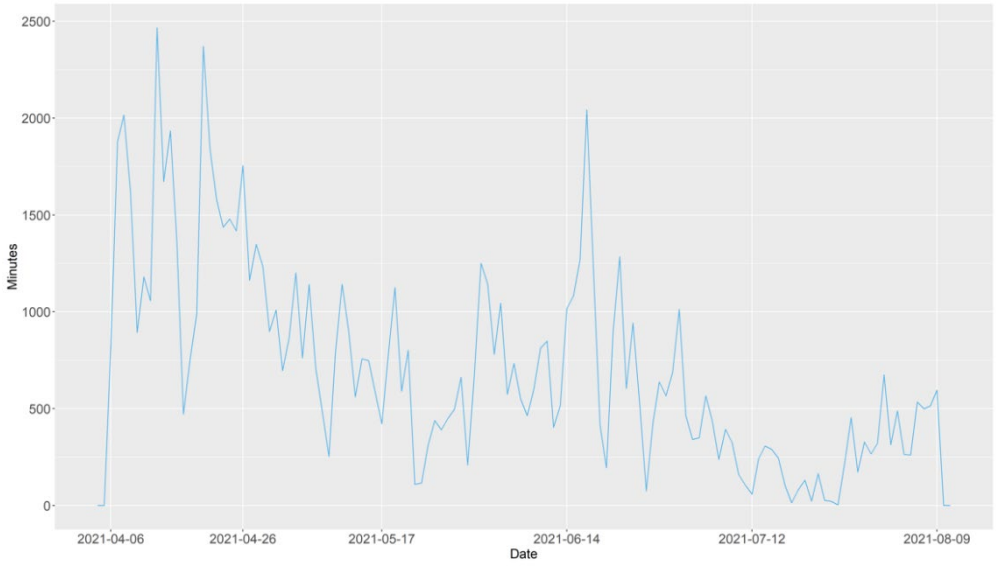

Figure 4. Lecture screencasts – Streamed minutes per day.

The data shows that the average number of streams per student and chapter of the lecture screencasts is always, and sometimes considerably so, larger than the number of videos (e.g. by approx. 55 percent for the *R* introduction), indicating that students take advantage of the option to rewatch contents and/or to break the contents into smaller units. A look at the *time factors* per chapter displayed in Figure 5, defined as the ratio of the total streaming minutes in that chapter to the total duration of its videos per student,

indicates that the repeatability indeed plays a major role in the use of the screencasts. The median *time factors* for the lecture screencasts ranging from 1.01 to 1.06, the maxima from 1.25 to 3.36.

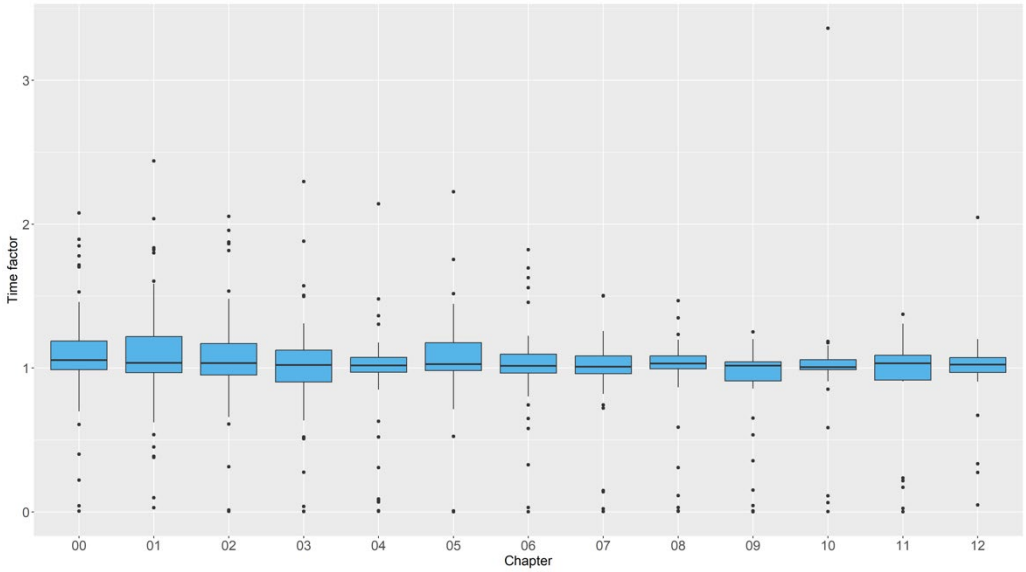

Figure 5. Lecture screencasts – Time factors.

The time autonomy offered by the screencasts is used by the students, as the timing throughout the day of the streams of the lecture screencasts in Figure 6 illustrates. While most streaming activity is concentrated between 10:00 a.m. and 06:00 p.m., the time span from the end of the lecture period through the e-exam (period 5) shows a stronger tendency towards *nightshifts*. The same can be seen for the e-tutorials.

ANALYSIS OF THE E-TUTORIAL DATA

The participation rates for the e-tutorials are drastically lower than those for the lecture screencasts. While it is approx. 80 percent for the first chapter for the latter in the complete student group, it is only approx. 55 percent for the former. While the actual completion rates for the lecture screencasts are high, they are considerably lower for some of the e-tutorials, as can be seen in Figure 7. The median of the share of the maximum number of points reached is lowest in the final e-tutorial: approx. 61 percent.

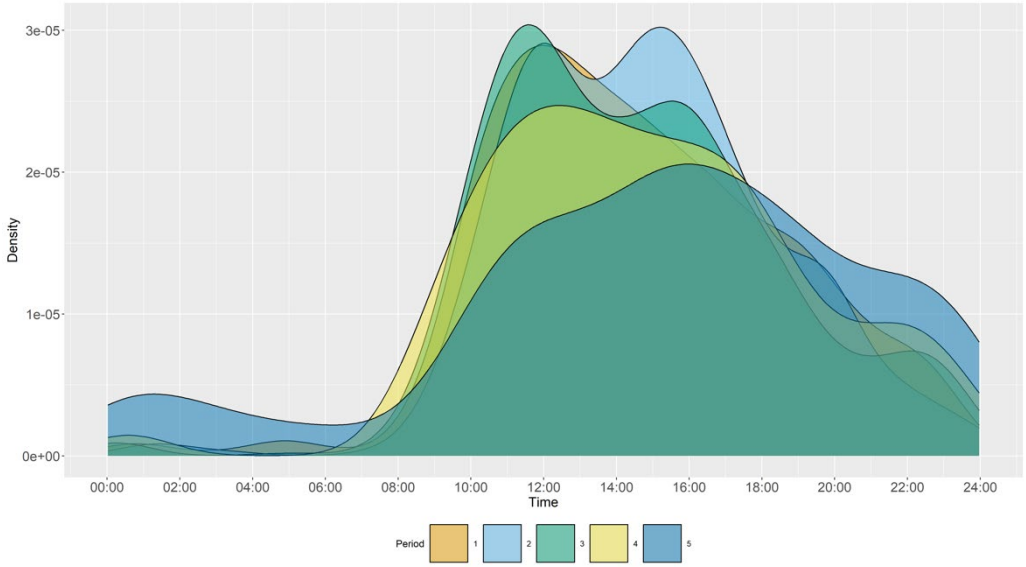

Figure 6. Lecture screencasts – Distribution of times of day of streaming.

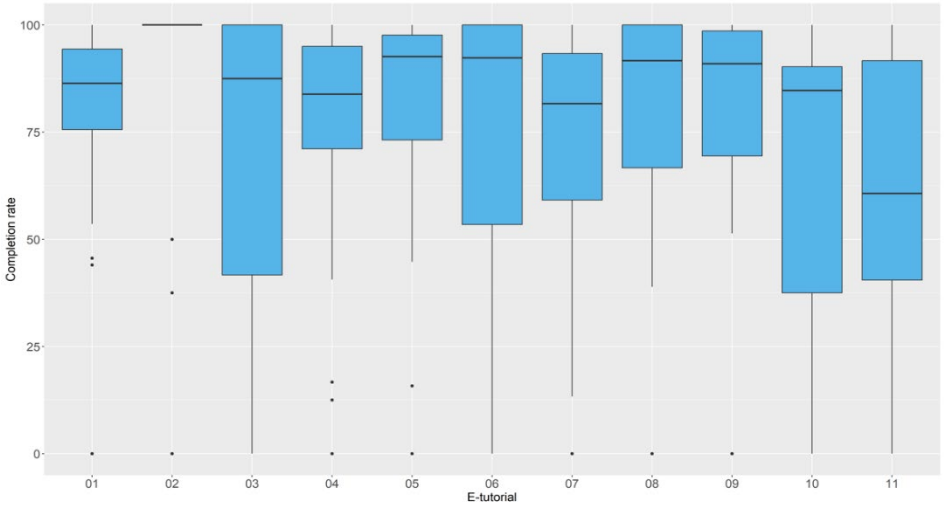

Figure 7. E-tutorials – Completion rates.

Although the median share of complete passes of an e-tutorial (meaning solving all subproblems in one pass) can be as low as approx. one fifth, the option of additional passes is clearly used by students, with median numbers of passes ranging from 2 to 5, and maxima as high as 39. As would be expected and is shown in Figure 8, in most cases students' performance actually improves after a non-perfect first pass if they start at least one additional pass.

USE OF DIGITAL COMPONENTS AND EXAM PERFORMANCE

The correlations of the completion rate in the e-exam to the completion rates in the three digital components of the course vary: from approx. 0.12 (tutorial screencasts), through 0.43 (lecture screencasts), to 0.72 (e-tutorials). If we were able to control for a proxy of, e.g., mathematical abilities, we could get a clearer picture of the influence of exercises on exam performance. Given our findings, it still seems safe to assume that active work with data (e-tutorials) is much more important than a mere passive consumption of solution strategies (traditional tutorials). Our own teaching experience leads us to believe that at least some students do not study consistently throughout the lecture period but rather delay studying, at times to an extreme degree. That this is detrimental to exam performance can be seen in Figure 9 that shows a time series of the cumulative completion rates of the lecture screencasts. Blue lines refer to students that passed the exam, orange ones to students that failed it. The thicker lines mark the average completion rates of the two groups. The pattern is similar for the e-tutorials.

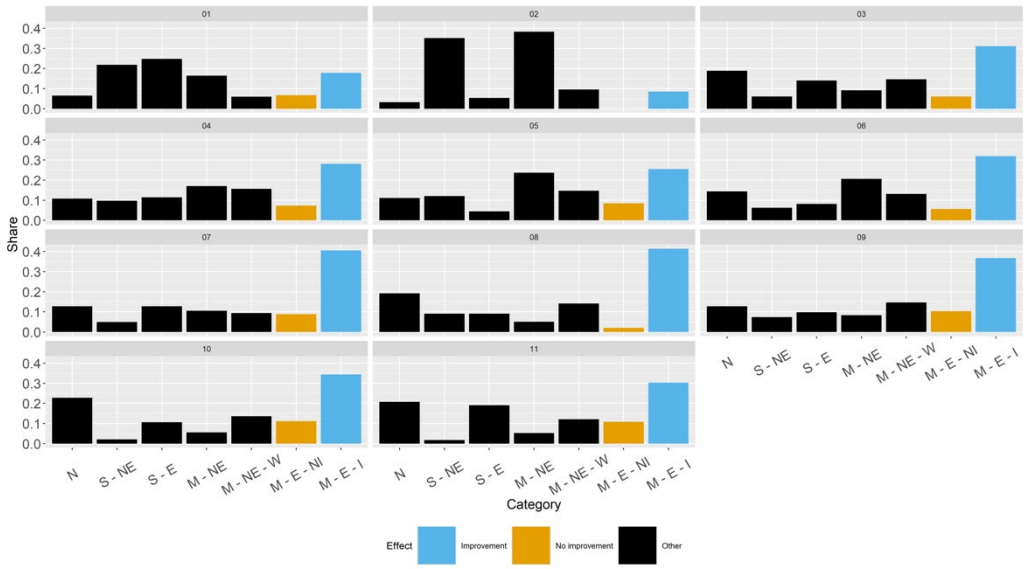

Figure 8. E-tutorials – Exercise indicator.

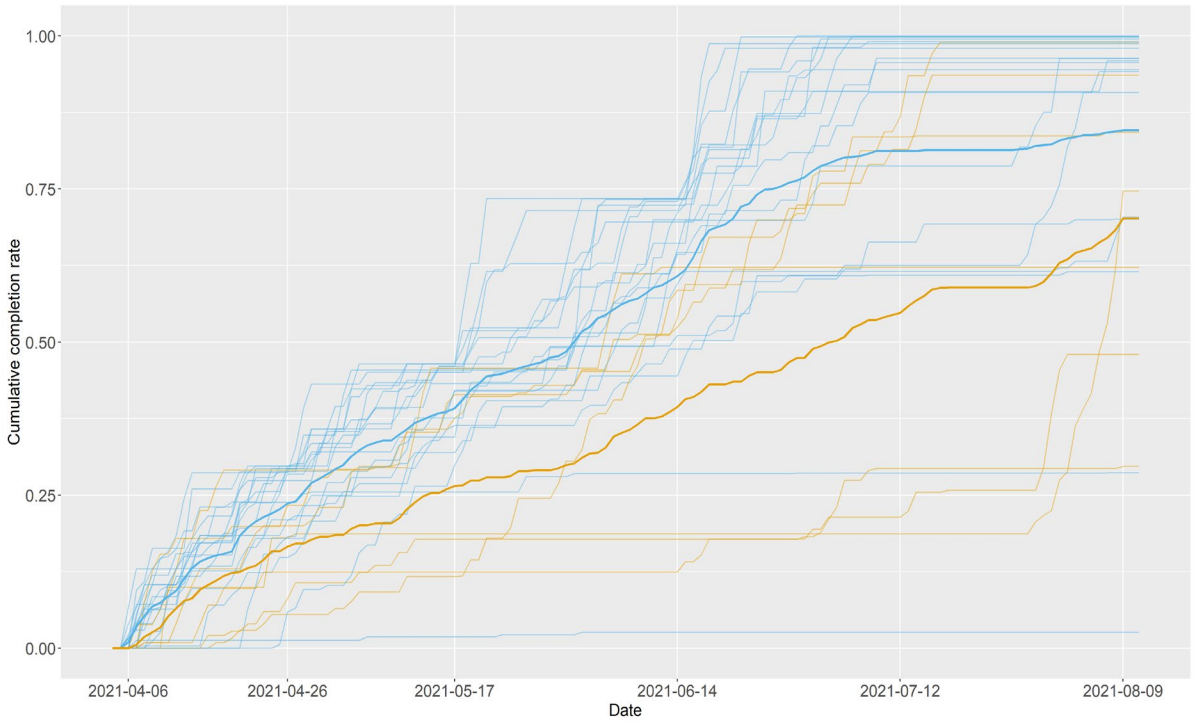

Figure 9. Lecture screencasts – Cumulative completion rates

CONCLUSION

We were able to track the use of digital components in a group of students in an undergraduate statistics course. The empirical findings mirror the intuition. Students make ample use of the flexibility offered by the 24/7 availability of digital tools. This is especially crucial for the possibility of repetition of more demanding parts of the contents. When it comes to statistics education, the hands-on experience of actively and repeatedly working with data sets is important. While this is the case, we also see that participation in the e-tutorials is comparatively low, some students could show more tenacity in their use of e-tutorials, and that there are some considerable delays in study efforts. To improve this, a way forward might be the implementation of adaptive e-learning scenarios that increase motivation and prevent frustration. However, in order to be able to properly implement such a system, a deeper understanding of the mechanics involved is necessary. As the project DLISA is moving on, we are striving for a larger sample size. In summary, our data show that there is a clear case for e-tutorials, for the provision of learning videos (as a supplement to a traditional lecture), and perhaps even for a clearly structured flipped classroom approach.

REFERENCES

Höfler-Hoang, B., Röder, D., & Ertz, F. (2020). Digitalisierung als Teil der Universitätsentwicklung – Strukturen, Angebote und Ziele der Universität Trier. In R. Stang & A. Becker (Eds.), Zukunft Lernwelt Hochschule – Perspektiven und Optionen für eine Neuausrichtung (pp. 35-43). de Gruyter. https://doi.org/10.1515/9783110653663-202

Münnich, R., Burgard, J., Höfler-Hoang, B., Nicknig, J. & Zimmermann, T. (2011). Individualisiertes eLearning – Eine innovative Anwendung auf die statistische Grundausbildung an der Universität Trier. Hamburger eLearning Magazin, 7, 51-52.