B.2 Limitations in Written Summative E-Assessment in Higher Education – An Analysis of a Student Survey

Anne Jantos, Charlotte Jung, Alexander Kohl
Technische Universität Dresden,
Professor für Wirtschaftsinformatik insb. Informationsmanagement

1 Introduction

Written summative online examinations are usually conducted virtually from remote locations (Bloh, 2006) and offer various advantages and challenges like high flexibility, low travelling cost and lower climate impact due to less paper consumption (Alruwais et al., 2018; Guàrdia et al., 2017). But virtual methods will not necessarily simplify the examination process at universities (Broadfoot, 2016). Observations at Technische Universität Dresden (TUD) have shown that even with a high level of effort in creating summative e-assessment online, it is hardly possible to develop a widely accepted method for implementation of written online exams mostly because it is technically complicated and leaves not enough room for various didactical approaches.

Summative e-assessment has been the exception before 2020 (Riedel & Möbius, 2018) and the rush to digitize written exams due to the pandemic leaves both students and teachers dissatisfied with the outcome of the many written online exam approaches (Handke & Schäfer, 2012). Research shows how socio-demographics influence the success of e-assessment (Bahar & Asi, 2018) or address security issues for users (Uotinen et al., 2020). But there is no research so far on specific technical limitations that influence students' performance in written online exams. This paper addresses that gap with a quantitative analysis of a survey of business and economics students at TUD in the winter semester 2020/2021, who were examined exclusively virtually due to the pandemic. With these findings, new technical and didactical methods for the implementation of summative e-assessment can be developed.

2 Theoretical background

Testing methods such as summative assessment summarize a student's achievement and are designed for reporting the student's performance at the end of a course of study, especially for certificate purposes (Black, 1986). It is a passive method that has no immediate effect on learning (Black, 1986; Challis, 2005) such as changing their field of study or dropping out of university altogether. Implementations of summative e-assessment are virtual written or oral exams that take place at the close of a learning experience with students and the examiner being in different locations (Challis, 2005; Bloh, 2006).

The online exams in the winter semester 2020/2021 were conducted using MS Forms and OPAL EXAM – platforms that are used for e-assessment at the TUD.

Both offer quizzes and free text as well as matching tasks and the possibility to mix questions and answer options and to select tasks from pools to achieve a higher variance of the tasks. No proctoring or other anti-cheating methods were applied other than mixing questions, and answers in single-choice questions and disable the possibility to jump between questions.

3 Methodology

This paper presents a deductive research approach based on primary data, which was analysed quantitatively and qualitatively. The data were collected once via an online survey using MS Forms. The sampling frame includes students who took online exams at the Faculty of Business and Economics at TUD in the winter semester 2020/2021 with a total of 2,294 students. 337 participants responded to the questionnaire and 183 participants described their technical difficulties in an open question.

3.1 Research proposition

Research on stress in written e-assessment was conducted before the pandemic. A survey showed that students do not experience significant stress from the online format and the corresponding technical measures of a written exam and therefore feel they performed equally well in analogue and virtual formats (Dermo, 2009; Rolim & Isaias, 2019; Okada et al., 2019). The following research proposition focuses on this aspect under the current pandemic conditions regarding technical issues. *RP1: Do students feel they perform worse in written online exams in winter semester 2020/2021 due to technical problems?* To investigate the students' technical problems in more detail (Appiah & Van Tonder, 2019) and to develop an explicit recommendation for reducing technical hurdles in the future, the following research proposition is posed. *RP2: Which limitations cause students to perform worse in written online exams?*

3.2 Data collection and analysis

Based on the research propositions a questionnaire was created to collect the data. It consists of eleven sections with a total of 48 questions to collect empirical data. A pre-test was conducted with four people to verify comprehension and clarity of the questionnaire items. The survey was accessible 4 weeks after the exam period in winter semester 2020/2021 for 17 days. Only the socio-demographic data and the stated course of study are considered in this paper with regard to the following two questions.

This paper analyses the responses to the statement "I would have done much better in an analogue examination with the same questions because of the technical problems" and the corresponding statements in text form. There were no explanations about what constitutes a technical problem to get the answers as uncorrupted as possible. The possible answers were "strongly agree", "mostly agree", "partly agree/partly disagree", "mostly disagree", "strongly disagree" and "no specification".

The data were analysed using qualitative content analysis according to Mayring & Fenzl, (2019). Firstly, the responses were reviewed to gain an overview. Then, the first 30% of responses were assigned to categories by a team of researchers. In the second step, the following 70% of responses were added to the existing categories by the researchers individually. Figure 1 shows the agreed-upon categories and the number of answers that were sorted into each category.

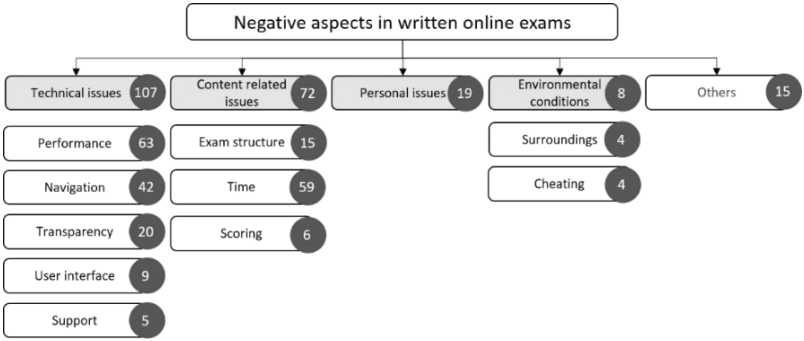

Figure 1: Category overview

Technical issues are the focus of this paper and concern all software and hardware problems, as well as technical support, including:

- Performance issues: Problems regarding loading times, internet problems and accessibility of the platform
- Navigation issues: Questions cannot be completed in an individual order and questions that have already been answered cannot be re-edited.
- Transparency: Overview of the type of questions and tasks, time frame and scoring of the entire exam
- User interface: The user interface of the exam that was provided
- Support: Help with problems before, during and after the exam by the examiner

Content related issues are problems related to the exam structure, time problems and time pressure and scoring. Personal issues include all individual influences, such as stress and nervousness. Environmental conditions include all external influences during the test and possibilities of deception such as surrounding (location of the examination and associated external influences) and cheating opportunities. In the category “Others” statements that could not have been sorted into the existing categories were combined. Although both questions were explicitly designed for technical aspects, the respondents often also gave answers that were aimed at organisational, didactical or environmental factors.

4 Results

337 responses were received in total. 183 participants explained their response to the statement: "I would have scored much better in an analogue exam with the same questions because of the technical problems". The written statements range from single words to long detailed answers. The average number of characters for free text answers is 278 which is surprisingly high for a voluntary answer.

4.1 Perceived exam performance

337 answers were analysed to address research proposition RP1. 9.1% of participants agree strongly that they would have performed much better in the usual format. Figure 2 shows these numbers in more detail.

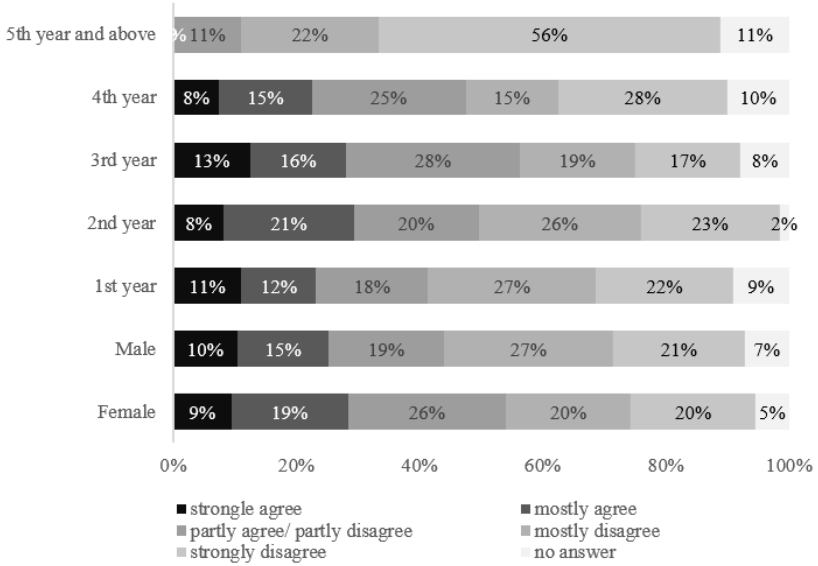

Figure 2: perceived disadvantage in written online exams due to technical problems regarding gender and years of study

16% agree mostly, 21.4% partly agree/partly disagree, 21.1% disagree mostly, 25.5% disagree completely. So, the majority (46.6%) of students feel they did not have a disadvantage during this year's written online exams due to technical problems. 25.1% feel the opposite. Male participants are slightly less inclined to perceive the online format brought a disadvantage than female participants. There is also a slight decrease with increasing studied semesters suggesting that with more experiences in higher education in general comes less inclination to feel that the written online exams bring disadvantages compared to the regular format in attendance.

Students in the 5th year and above especially do not feel any disadvantage due to technical problems. Figure 2 shows the results regarding gender and the number of studied semesters. 25% of business pedagogics, 27% of business informatics, 50% of business engineers and 55% of business economics stated they felt a disadvantage in written online exams due to technical difficulties.

So, the data suggest that the research proposition *RP1: "Do students feel they perform worse in written online exams in winter semester 2020/2021 due to technical problems?" cannot be generally answered* because it is an individual appreciation for the exam situation that affects the outcome. But 47% of all participants say they feel at least partly at a disadvantage because of technical problems in written online exams. There is a great variation between students of different fields of study and gender as well as progress in their studies.

4.2 Technical limitations in written online exams regarding socio-demographic details

To address *RP2: "Which limitations cause students to perform worse in written online exams?"* the variance of the different technical problems is analysed based on the qualitative evaluation of the open questions "Please explain your decision on the last question here (voluntary)" in regards to the statement: "I would have scored much better in an analogue exam with the same questions because of the technical problems". Figure 4 shows that the qualitative analysis is an amplification to the previously discusses analysis regarding gender and the field of study.

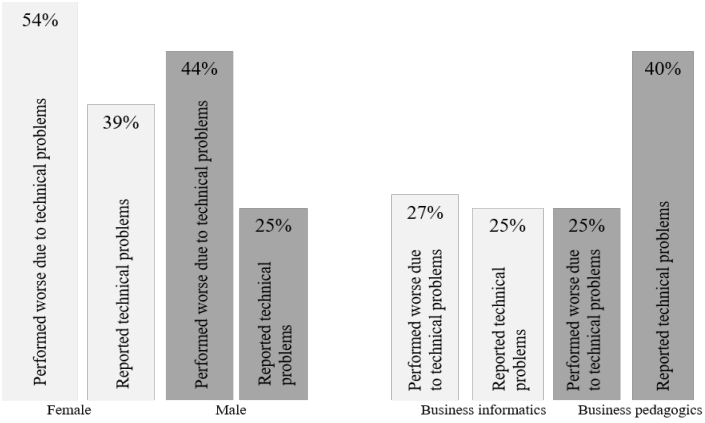

Figure 4: Results of the quantitative vs. qualitative analysis in contrast

44% of male and 54% of female students stated they found technical problems to worsen their exam outcome. 25% of male and 39% of female participants stated in their written answers to having technical difficulties. Regarding the field of study, a contradiction formed itself in the qualitative analysis. 25% of business pedagogics and 27% of business informatics stated they performed worse yet, 25% of business informatics and 40% of business pedagogics claimed in detail that they have had technical difficulties. The data concerning the progress in the students' studies is confirmed by the qualitative analysis. There is a considerable decrease in technical difficulties with increased years of study. 34% of students in their first and second year, 28% in their third and fourth but only 11% in their fifth year or above complain about technical difficulties.

4.3 Variations of technical limitations in written online exams

In the following, the categories of technical problems during written online exams formed are considered regarding their frequency, the combination with each other and distribution among the socio-demographic details. 107 students stated that they had experienced technical problems. Figure 3 shows the results:

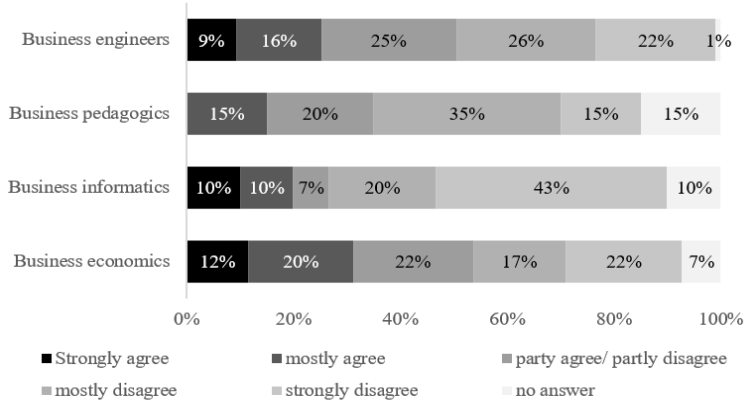

Figure 3: perceived disadvantage in written online exams due to technical problems regarding the field of study

Technical problems are subdivided into performance, navigation, transparency, user interface and support. For this analysis, only those participants who answered "mostly agree" and "strongly agree" to feeling they performed worse due to technical issues are analysed. Participants who stated that they "partly agree/partly disagree" to having performed worse will not contribute to this analysis.

With 63 mentions, the subcategory "*performance* problems" is the category that shows the strongest expression. In contrast, the problem of *transparency* was only mentioned 19 times. If this were to be inferred, performance problems would be the greatest cause of poor performance in online exams. However, only 49% of the respondents who described performance problems performed worse in the online exams (table 1). In contrast, 63% of respondents who reported problems in transparency performed worse (table 1). It can thus be deduced that problems in transparency led more frequently to a worse exam result than performance problems. Thus, the subcategory transparency is the subcategory with the greatest impact on the exam result. It can be assumed that performance problems occurred frequently, but that these were compensated for by additional time or support from the examination supervisors. Table 1 shows the categories and their influence on perceived exam performance.

Table 1: Categories and their influence on perceived exam performance

Category	**Mentions**	**Percentage of participants** who named aspects of this category and felt they performed worse in the online exam due to technical difficulties.
Technical	107	51 %
Performance	63	49 %
Navigation	42	57 %
Support	5	40 %
Transparency	19	63 %
User Interface	8	50 %
Content	72	43 %
Exam Structure	15	47 %
Time	59	44 %
Scoring	6	33 %
Environmental Issues	8	38%
Surroundings	4	75 %
Cheating	4	25 %
Personal	19	58 %

In contrast, the problem of transparency occurs mainly due to a deliberately selected linear examination format and is not additionally compensated for with other measures like an overview over the questions and tasks and the corresponding scoring, which would help students to orientate themselves in advance and plan their priorities. However, it must also be considered that these linear examination formats prevent cheating attempts, especially when combined with a random question order.

This prevents students from taking exams together. To prevent cheating attempts and still create a clear exam, students could be informed at the beginning of the exam about the thematic content of the entire exam and how the points and thus the time required are distributed among these tasks. 57% of students reported having experienced *navigation* problems. 50% of the students reported problems with the *user interface* experienced. 40% complained about problems with *support***.** Participation in the mock exam only slightly reduces the probability of encountering technical problems. 25% of the examinees who took the mock exam had technical problems. Of those who did not take part, 29% had technical problems. This leads to the answer to the second research *RP2: "Which limitations cause students to perform worse in written online exams?"*. In terms of technical issues *transparency* and a clear overview of the entire examination is the decisive factor for the satisfactory completion of an examination. If it is disturbed, students find it very obstructive and therefore perform worse.

4.4 Problems in written online exams other than technical limitations

The category "technical issues" was selected most frequently (107 responses), followed by the category "Content issues" (72 responses) and the categories "Personal issues" (19 responses) and "Environmental Conditions" (8 responses). This means that in addition to technical problems, students also experienced problems in other areas and these also led to a poorer result in online examinations.

Content limitations include structure issues, timing issues and scoring issues. The structure was mentioned negatively 15 times. Students complained about unusual formats and irregular testing methods. Scoring issues were mentioned 6 times. In some of the arithmetic tasks, points were only awarded for the overall result, but not for the arithmetic path, which was perceived as unfair. The most important category is *time*. 59 students mentioned that time factors have hurt their performance. The upload of individual work results sometimes took a considerable time and the duration was not added to the examination period, which was described as unfair. Furthermore, students stated that the extent of the tasks was too large for the corresponding time. 44% of those who described difficulties in the category time felt they performed worse in the exam (table 1).

Personal issues were mentioned 19 times. These include factors such as stress and nervousness (Dermo, 2009). Students explained that uncertainties regarding technical difficulties made them feel nervous about the exam since they were not assured what would happen in case they could not attend due to technical problems. Respondents indicated, in the category *environmental issues,* that the opportunities for cheating created by the online format hurt their motivation, as it was likely that many would cheat to gain an advantage, which was confirmed by a related survey (Jantos, 2021). Another negative aspect was the lack of an exam atmosphere, as the respondents reported disruptive factors in their workplace and surroundings.

4.5 Recommendations

The category with the greatest influence on the perceived performance in written online examinations is *transparency*. That means that students criticise the lack of an overview in the process of the exam and describe it as disturbing. To avoid this problem, the approach of making the exam available in its entirety and thus allowing the student to see all the tasks before he or she starts working on them is suitable. However, this leads to a higher probability of cheating. Another way is to show the students the procedure transparently in preparation for the exam without going too much into the content. This can enable the student to skip tasks and prioritise them sensibly without increasing the possibility of cheating.

Time is the second most relevant category. Students report a general shortage of time to complete assignments, but also address the fact that time lost due to upload processes is not compensated for or other technical problems lead to delays and time loss is not automatically credited back. This can be addressed by planning a generally larger time frame or smaller task extent. However, time constraints were also perceived as a problem before the switch to the virtual format and it is sometimes part of the examination strategy to stress students' time to test their performance. Other alternatives to written summative examinations should also be considered. For smaller numbers of participants, an oral examination is strongly preferred. Here, technical problems can be unilaterally coordinated and addressed accordingly without putting the student in a worse position. Furthermore, the possibilities of cheating are significantly lower in this format (Jantos, 2020; Kaiiali et al., 2016).

Online teaching leads to a change in the learning culture. Technical problems are used as an excuse to cover up a lack of preparation. This can sometimes also be observed in examination situations. Here the lecturer should communicate his expectations of the students' working methods early on to address possible problems in the online formats during the course and to practise handling them so that technical problems no longer influence the performance during the exam. Furthermore, the use of online formats for examinations should be made more a part of everyday study. However, this requires an organisational and pedagogical turnaround in university teaching, especially considering that all students in the faculty are now digital natives and therefore also have corresponding needs. Technical socialisation in the curriculum is long overdue. (Schiefner et al., 2020; Prendes et al., 2021)

Finally, it should be noted that it is not the purely technical aspects of a summative online examination that students mainly find obstructive, but the didactic and organisational ones. The simplest solution to the main problem of transparency is to improve coordination between the teacher and the exam takers.

5 Discussion

Summative e-assessment in winter semester 2020/2021 at TUD was a success in terms of implementation since all exams were conducted and no student had to wait for a year to be examined and neither students nor faculty were forced to attend an exam and be exposed to harm. This qualitative analysis has shown that students do not unanimously perceive a virtual format for written exams as negative, disruptive, or hindering. Almost half of the respondents affirmed that their performance was hindered because of technical problems. However, the other half denied this. According to this analysis, the reason for the disruption of the examination performance is mainly a lack of transparency in the presentation of the examination and too little time given to work on the examinations. Performance problems were the most frequently mentioned category but had less influence on the students' performance than transparency.

Since the results of this analysis give only a glimpse into the extensive situation of higher education e-assessment the results cannot be scaled up to all students globally. Only one semester was surveyed, no long-term conclusions can be drawn especially since the exams were taken under the conditions of the pandemic and it can be assumed that approaches to virtualizing exams were rushed. Because concepts were implemented prematurely or hastily, the dissatisfaction cannot be applied generally to all written e-assessment concepts. Furthermore, the results cannot be transferred to e-assessment in attendance, because here the organisational installation of hurdles to prevent cheating can be dispensed with. And it can be assumed that many technical hurdles that are linked to the students' infrastructures are prevented in an e-assessment in a controlled examination environment. Also, the sample is a comparatively homogeneous group because it consists almost exclusively of German students studying economics, but the study represents the entire faculty of business and economics of TUD, so, it is still possible to form valid conclusions.

So far, only the students' point of view was considered here, so a survey of the faculty's lecturers is the logical next step. A qualitative study of the staff of the faculty will be implemented soon to also explore this side of the technical hurdles of the e-assessment and to be able to make generally valid statements. Furthermore, it should be investigated which communication strategies are useful to support the process of a written online examination as transparently as possible and to take away the students' uncertainties by introducing the technology and explaining the process. In addition, the aspect remains open that comfortable handling of the examination software and the desire to avoid cheating are contrary to each other. This can only be achieved if there is mutual trust, which again requires meaningful communication between teacher and student. This analysis of the student survey shows that virtual formats of summative exams have many limitations. New approaches for the future must learn from the circumstances of the pandemic and take on the advances it has brought to approach assessment in a sensitive and learner-centred way (Schoop et al., 2021; Altmann et al., 2019, Altmann et al., 2021).

Literature

Alruwais, N., Wills, G., Wald, M. (2018). Advantages and Challenges of Using e-Assessment. International Journal of Information and Education Technology, 8(1), 34–37. https://doi.org/10.18178/ijiet.2018.8.1.1008

Altmann, M., Clauss, A., Jantos, A., Lenk, F., Reeb, S., Safavi, A. A., & Schoop, E. (2019). Digitalisation in higher education: A flipped classroom arrangement to foster internationalization. In. 22. Workshop GeNeMe'19 Gemeinschaften in Neuen Medien (2019) (pp. 127–130)

Altmann, M, Langesee, L.M., Misterek, J. (2021). Designing Formative Feedback Guidelines in Virtual Group Work from frim a Student's Perspective. In 13th International Conference on Education and New Learning Technologies

Appiah, M.; Van Tonder, F. (2019). Students' Perceptions of E-assessment at a Higher Education Institution. In *5th International Conference on Computing Engineering and Design (ICCED).* pp. 1–7, https://doi.org/10.1109/ICCED46541.2019.9161088.

Bahar, M., Asi, M. (2018). Attitude towards e-assessment: influence of gender, computer usage and level of education, Open Learning: The Journal of Open, Distance and e-Learning, 33:3, 221–237

Black, H. D. (1986). Assessment for learning. In D. L. Nuttall (Ed.), Assessing educational achievement. London: Falmer.

Bloh, E. (2006). Methodische Formen des E-/Online-Assessment unveröffentlichtes Manuskript, Kaiserslautern.

Broadfoot, P. (2016). Assessment for Twenty-First-Century Learning: The Challenges Ahead. Learning, Design, and Technology. Springer, Cham, 1–23.

Challis, D. (2005). Committing to quality learning through adaptive online assessment Assessment in Education 30(5): 519–527.

Dermo, J. (2009). e-Assessment and the student learning experience: A survey of student perceptions of e-assessment. In British Journal of Educational Technology. doi:10.1111/j.1467-8535.2008.00915.x

Guàrdia, L., Crisp, G., Alsina, I. (2017). Trends and Challenges of E-Assessment to Enhance Student Learning in Higher Education. In J. Keengwe, E. Cano, & G. Ion (Eds.), Advances in Higher Education and Professional Development. Innovative Practices for Higher Education Assessment and Measurement (pp. 36–56). IGI Global. https://doi.org/10.4018/978-1-5225-0531-0.ch003

Handke, J., Schäfer, A.M. (2012). E-Learning, E-Teaching und E-Assessment in der Hochschullehre: Eine Anleitung, Oldenbourg Verlag

Jantos, A. (2020). Conducting Oral Interviews Virtually using MS Teams. In 23. Workshop GeNeMe'2020 Gemeinschaften Neue Medien. P. 294–298

Jantos, A. (2021). Motives for Cheating in Summative E-Assessment in Higher Education – A Quantitative Analysis. 13th International Conference on Education and New Learning Technologies.

Kaiiali, M., Ozkaya, A., Altun, H., Haddad, H. , Alier, M. (2016). Designing a Secure Exam Management System (SEMS) for M-Learning Environments. In IEEE Trans. Learning Technol. 9 (3). P. 258–271. https://doi.org/10.1109/TLT.2016.2524570

Landis, J.R., Koch, G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33 1, 159–74

Mayring, P., Frenzel, T. (2019). Qualitative Inhaltsanalyse

Okada, A., Noguera, I., Alexieva, L., Rozeva, A., Kocdar, S., Brouns, F., Ladonlahti, T., Whitelock, D., & Guerrero-Roldán, A–E. (2019). Pedagogical approaches for e-assessment with authentication and authorship verification in Higher Education. British Journal of Educational Technology, 50(6), 3264-3282. https://doi.org/10.1111/bjet.12733

Prendes-Espinosa M.P., Gutiérrez-Porlán I., García-Tudela P.A. (2021). Collaborative Work in Higher Education: Tools and Strategies to Implement the E-Assessment. In: Workgroups eAssessment: Planning, Implementing and Analysing Frameworks. Intelligent Systems Reference Library, vol 199. Springer, Singapore. https://doi.org/10.1007/978-981-15-9908-8_3

Riedel, J., Möbius, K. (2018). Bestandsaufnahme, Hindernisse und Möglichkeiten des Einsatzes von E-Assessment an sächsischen Hochschulen. Beiträge Zur Hochschulforschung, 40. Jahrgang, 4/2018.

Rolim, C., Isaias, P. (2019). Explaining the use of e-assessment in higher education: Teachers and students' viewpoints. In British Journal of Educational Technology. Vol 50 No 4. P. 1785–1800. https://doi.org/10.1111/bjet.12669

Schiefner-Rohs, M., Hofhues, S., Aßmann, S., Brahm, T. (2020). Studieren im digitalen Zeitalter. Methodologische Fragen und ein empirischer Zugriff – In Beiträge zum 26. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft. S. 337–348. urn:nbn:de:0111-pedocs-192525

Schoop, E., Sonntag, R., Altmann, M. & Sattler, W. (2021). Imagine it's "Corona" – and no one has noticed. Lessons Learned: Spin Offs of Digital Teaching Experiences Vol. 1, No.1&2.

Uotinen, S., Ladonlahti, T., Laamanen, M.(2020). Developing E-Authentication for E-Assessment – Diversity of Students Testing the System in Higher Education. European Journal of Open, Distance and E-Learning,23(2) 99–115. https://doi.org/10.2478/eurodl-2020-0013