Comparison of 2D Virtual Learning Environments with Classic Video Conferencing Systems for Tertiary Education

Gerhard Hube, Kevin Pfeffel, Nicholas H. Müller
Faculty of Economics and Business Administration and
Faculty of Computer Science and Business Information Systems
Technical University of Applied Sciences Würzburg-Schweinfurt
Würzburg, Germany
e-mail: gerhard.hube@thws.de
e-mail: kevin.pfeffel@thws.de
e-mail: nicholas.mueller@thws.de

***Abstract*— This work is a follow-up to our previous study "2D Virtual Learning Environments for Tertiary Education", which was carried out in 2022. The main focus was to analyze the suitability of a 2D Virtual Learning Environment (VLE) for tertiary education using the desktop based 2D immersive environment 'gather.town'. The study was conducted with a selected course of a Master's program at the Technical University of Applied Sciences Würzburg-Schweinfurt over one semester. Accompanying the course, subjects were asked to complete the Online Learning Environment Survey (OLLES) questionnaire weekly for analysis, and additional qualitative interviews were conducted afterwards. The descriptive analysis suggests that the immersive 2D environment used is holistically suitable as a learning environment in the tertiary sector, due to high and very high values for presence, participation, collaboration and active learning. For this paper, two seminars were conducted using Virtual Learning Environments, one of them in 'gather.town' and the other in 'Zoom'. In addition to the OLLES questionnaire and the qualitative interviews, the Igroup Presence Questionnaire (IPQ) was also queried. Additionally, the exam grades were also collected as a performance measure. This made it possible to compare the different learning environments. When comparing the questionnaires, only some dimensions showed a difference between Virtual 2D Learning Environments and Classic Video Conferencing Systems. In contrast, with exam grades, subjects were found to perform better with Virtual 2D Learning Environments than with Classic Video Conferencing Systems.**

Keywords-Virtual Learning Environments; Online Teaching; Tertiary Education; 2D Environments; Desktop Virtual Reality; Zoom; gather.town.

I. INTRODUCTION

This contribution is based on the first step of the study published in 2022 in the International Journal on Advances in Systems and Measurements, vol. 15, no. 3 & 4 with the title "2D Virtual Learning Environments for Tertiary Education" [1]. As the main result of the study, the high scores of the OLLES [2] questionnaire can be mentioned. In connection with the interviews, it can be said that an Immersive 2D Environment can be used holistically as a form of teaching and has advantages over Classic Video Transmission Tools. As a practical implication, it can be deduced that the use of Virtual Learning Environments in the tertiary sector, on the one hand, can be relatively easily deployed with existing software solutions and, on the other hand, are also well received and therefore offer benefits for students.

Nevertheless, this first study was only an overview of the use of an immersive 2D environment as a learning tool within tertiary education. Group comparisons with other teaching formats were not possible. Therefore, this is the goal for this research. Here, the same teaching unit is being tested again in gather.town and at the same time another teaching unit is being tested in Zoom. Again, the OLLES questionnaire is used and additionally the Igroup Presence Questionnaire (IPQ) [3]. The IPQ is a scale for measuring the sense of presence experienced in a Virtual Environment (VE). The qualitative interviews were also be used again for data collection. In Figure 1, there is an overview of the timeline and the different learning environments and measuring instruments.

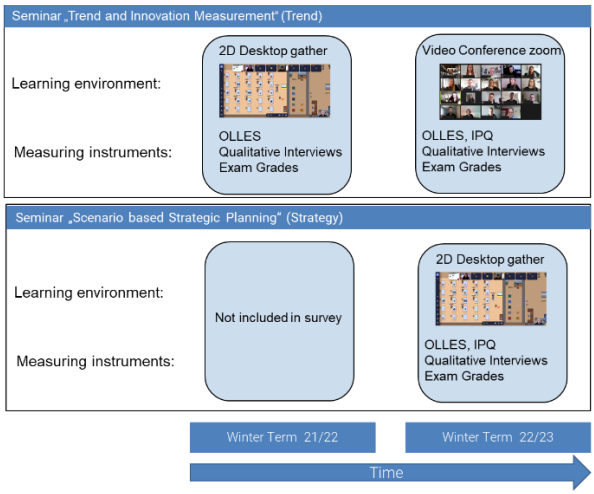

Figure 1. Overview of timeline, seminars, learning environments and measuring instruments for the study.

With the results of the different seminars and learnings environments, a comparison of the two forms of teaching can be made.

 ISBN: 978-1-68558-089-6

The definitions and explanations for the basic terms, as well as the overview of studies and related works about Virtual Learning Environments (VLE) and Virtual Reality (VR) in higher education were made in the study from 2022 [1].

Additional, to the literature review from our study in 2022 [1], there were several new studies published about educational online learning, especially with Learning Management Systems (LMS) like Moodle and Video Conference Systems, especially Zoom [4] - [8]. In addition, many studies about the phenomenon of "Zoom fatigue" were published [9] - [13] which underlines the need for alternative online Learning Environments like low immersive Desktop Environments. Probably because of this need, several studies appeared with gather.town as one example for this kind of Virtual Environment. Lo and Song [14] performed a review of the empirical studies in gather.town and revealed that there is still a lack in studies besides computer science courses, the examination of student's behavior and learning achievements. The authors also found out, that most of the studies had only a short duration and suggest studies with a longer duration. With this study, we evaluate Virtual Learning Environments over several semesters in the context of seminars, not in computer science, but in business administration. We also include exam grades for learning outcomes. With these conditions, we fulfill some of the requirements for further research. To summarize, so far we have looked descriptively at the suitability of 2D Virtual Learning Environments for tertiary education and now we want to test this statistically by means of a first comparison of 2D Virtual Learning Environments and Classic Video Conferencing Systems.

Following in Section 2, we explain the method used. Section 3 resumes the results, which are then discussed in detail in Section 4 with some limitations. Section 5 forms the end of the paper and contains the conclusion with the main results and future studies.

II. METHOD

In the following, we present the immersive learning environment gather.town, in which the course took place, and the measuring instruments OLLES and IPQ, which were used for the assessment. In addition, qualitative interviews were subsequently conducted with some of the subjects, which will also be presented here.

A. Immersive 2D environment gather.town

The software gather.town [15] was used as an immersive 2D environment. This is a web conferencing software, which allows to create a complete virtual replica of the teaching building. Within this virtual space, users can move around using avatars and interact with each other and their environment, similar to real life. If the avatars now walk around in the Virtual Environment and then meet each other at a certain distance, the camera and the microphone of the computers are automatically switched on, and the users have the opportunity to communicate. The graphical user interface is quite simple and it does not demand any special requirements to run on a variety of computers. In preparation, the entire real seminar building was recreated in the gather.town environment and the following Virtual Environment settings and software features were used:

The podium is the classic teaching situation, as shown in Figure 2. Within the gather.town environment, all students and the tutor are in one large room. The tutor stands in front at the lectern, while the students take their places at the tables. All students can see, hear and, of course, communicate with each other via camera and microphone. It is possible to share the screen to provide lecture slides or other content to all participants in the plenum area. In this way, the tutor can use lecture slides in addition to a verbal execution of the learning topic, as they would be used in a real teaching situation.

We refer to our publication in 2022 [1] for explanation of the features "Workshop", "Whiteboard", "Break Rooms" with games and yoga room, and "Interactive elements".

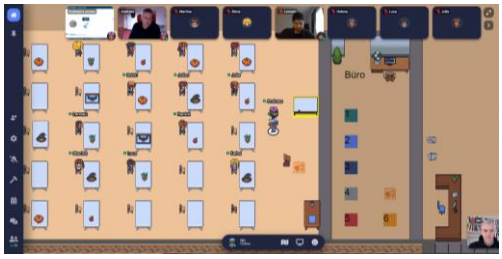

Figure 2. This is the podium. A classic teaching situation in a shared space is shown.

B. Video conference tool Zoom

Zoom is one of the Classic Video Conferencing Tools with quite wide spread usage for education, especially during the COVID-19 pandemic, but also after reopening universities in 2021 [4] [16]. With Zoom, it is possible for one or more people to interact through chat messages, video based visual communication, and group work [17]. Besides the communication in the whole group of participants, it is also possible to create subgroups (Break out rooms) for group work or group discussions. There is also the possibility to share the screen with other participants, to do little surveys and to use a whiteboard. The classic appearance is the monitor full of video tiles with the participants of the Zoom meeting, as shown in Figure 3.

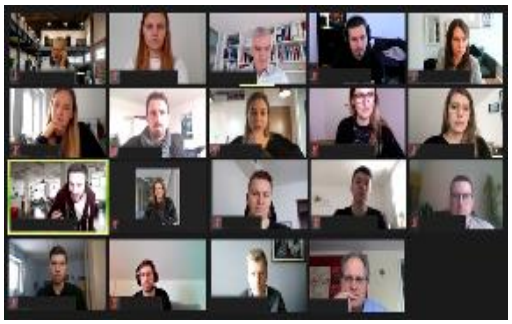

Figure 3. Video tiles on monitor while classical Zoom video conference.

 ISBN: 978-1-68558-089-6

C. Measuring instrument

The OLLES questionnaire in its modified 35-item form was used as the measurement instrument [2]. The OLLES questionnaire is a web-based survey instrument for use in online learning environments in tertiary education. In this context, the OLLES questionnaire provides inferences about students' perceptions of interaction opportunities within an online environment in terms of economy and efficiency. The dimensions of the OLLES are Student Collaboration (SC), Computer Competence (CC), Active Learning (AL), Tutor Support (TS), Information Design and Appeal (IDA), Material Environment (ME), and Reflective Thinking (RT). In addition, questions about general computer use and Internet use were also recorded. All items were measured on a 5-point Likert scale [18].

The IPQ [3] was also used. The IPQ is a scale for measuring the sense of presence experienced in a Virtual Environment. Here, the sense of presence is understood as the subjective sense of being in a Virtual Environment. Also, the igroup.org project consortium states that: "the sense of presence can be separated from the ability of a technology to immerse a user. While this immersion is a variable of the technology and can be described objectively, presence is a variable of a user's experience. Therefore, we obtain measures of the sense of presence from subjective rating scales." The IPQ has three subscales and one additional general item not belonging to a subscale. The three subscales are Spatial Presence (the sense of being physically present in the VE), Involvement (measuring the attention devoted to the VE and the involvement experienced) and Experienced Realism (measuring the subjective experience of realism in the VE). There is also a general item that assesses the general "sense of being there". This item has high loadings on all three factors, with an especially strong loading on Spatial Presence. The original questionnaire was constructed in German, so we used this one, since the subjects are German native speakers. All items were measured on a 7-point Likert scale with a range from 0 to 6 [18].

For the qualitative interviews, a separate questionnaire was developed, which can be viewed in full in our previous paper [1] where the same questionnaire was used. First, an introductory question was asked in order to lead the test persons into the interview situation in a relaxed manner and to check whether they could still remember the seminar well within the Virtual Learning Environment. Building on this, at least one question was asked about each dimension of the OLLES to develop a deeper understanding of why one of the dimensions had performed well or poorly. In addition, the questions of the questionnaire still investigate whether the subjects prefer face-to-face classes, a Virtual Learning Environment such as gather.town or Classic Video Conferencing Software such as Zoom and why this is so. Finally, the questionnaire examines whether the Virtual Learning Environment gather.town was also used outside the actual seminar and, if so, for what other purposes. In addition, questions are asked about the highlights and shortcomings of the software used.

Furthermore, exam grades were collected as a form of performance measure.

D. Experimental procedure

Even before the first seminar, all test persons were familiarized with the gather.town environment resp. the Zoom environment. In particular, the basic functions were tested, so that everybody knows them and can use them independently. In addition, the OLLES questionnaire was introduced, since this was used in its original English language, but the test persons were not native English speakers.

Both seminars were held over 5 days each, with one teaching session starting in the early afternoon and lasting 5-6 school lessons each. Both seminars were held exclusively in their respective VE used. There were a total of two time measurement points, one after the first seminar and one after the last seminar. Both questionnaires were completed online directly after the seminar.

The qualitative interviews were collected a few days after the last seminar, but they were conducted within gather.town resp. Zoom. An appointment was made with a respondent within gather.town resp. Zoom, where the interview was conducted and the audio track was recorded. The audio track was then transcribed, analyzed and interpreted.

E. Sample

All data were collected at the Technical University of Applied Sciences Würzburg-Schweinfurt within the seminars "Scenario Based Strategic Planning" (from here just "Strategy") and "Trend Analysis and Innovation Measurement" (from here just "Trend") of the master study program "Integrated Innovation Management". The seminar "Strategy" was held in gather.town and the seminar "Trend" was held in Zoom, as shown in Figure 1.

For the seminar Strategy, a total of 19 subjects participated. However, only 16 subjects completed the questionnaires. This leaves n = 16 valid subjects for the final analysis. The average age of the subjects is 25.19 years, with a minimum of 22 years and a maximum of 33 years. Of the n = 16 subjects, 5 are female (31.3 %) and 11 are male (68,7 %). In addition, it must be noted that only 11 subjects could be used for the comparison of the two measurement points, since only these 11 subjects completely filled out the two questionnaires. For the remaining statistics, however, all 16 subjects can be used. In addition, for a performance comparison in the form of the scores, the scores of all 19 subjects of the seminar were used. Five randomly selected subjects were used for the qualitative interviews. Afterwards, it was checked to what extent the answers of the subjects overlapped or whether new insights could still be gained with further surveys, but a feeling of saturation set in. Therefore, n = 5 interviews were considered sufficient. Of the n = 5 subjects, 3 are female and 2 are male.

For the seminar Trend, a total of 19 subjects participated. However, only 17 subjects completed the questionnaires. This leaves n = 17 valid subjects for the final analysis. The average age of the subjects is 25.06 years, with a minimum of 22 years and a maximum of 33 years. Of the n = 17 subjects, 6 are female (35.3 %) and 11 are male (64.7 %). In addition, it must be noted that only 10 subjects could be used for the

 ISBN: 978-1-68558-089-6

comparison of the two measurement points, since only these 10 subjects completely filled out the two questionnaires. For the remaining statistics, however, all 17 subjects can be used. For a performance comparison in the form of the scores, the scores of all 19 subjects of the seminar were used.

Four randomly selected subjects were used for the qualitative interviews. Afterwards, it was checked to what extent the answers of the subjects overlapped or whether new insights could still be gained with further surveys, but a feeling of saturation set in. Therefore, n = 4 interviews were considered sufficient. Of the n = 4 subjects, 2 are female and 2 are male.

III. RESULTS

The results section is divided into different areas. First, there is a statistical part in which the time measurement points of the individual subjects are compared to see if there is a difference between the first time measurement point after the first seminar unit and the last time measurement point at the end of all seminar units. This is complemented by a purely descriptive part, in which the mean values of the OLLES and IPQ questionnaires are considered. Thereafter is the part in which the results of the qualitative interviews are presented. Both of these parts are again subdivided into the individual seminars. Lastly, there is a statistical part. In this part, first there are group comparisons related to the values of the OLLES and IPQ questionnaires. The data from our previous paper [1] will also be used. Finally, there is a group comparison of the exam grades as a performance measure.

A. Results for "Strategy" using gather.town

First, the Wilcoxon test will be used to examine whether there are differences in the OLLES test between the individual time measurement points and thus whether there was a change in the evaluation with regard to the repetition of the use of gather.town.

Two time measurement points were not available for all 16 subjects, therefore the following Wilcoxon test was only calculated with n = 11 complete subjects.

The Wilcoxon test showed that there was no difference between time measurement point 1 and time measurement point 2 regarding the OLLES questionnaire.

Next, using the Wilcoxon test will be used to examine whether there are differences in the IPQ test between the individual time measurement points and thus whether there was a change in the evaluation with regard to the repetition of the use of the gather.town environment.

There was a significant difference of the variable G (General Presence). The statistic test is z = -2.850 and the associated significance value is p = .002. Thus, the difference is significant: the central tendencies of the two time measurement points differ (Asymptotic Wilcoxon test: z = -2.85, p = .002, n = 11).

For the other scales, there was no significant difference between time measurement point 1 and time measurement point 2.

The next step is a descriptive analysis of the mean value variables of both time measurement points together. This also includes all measured values regardless of whether there were only one or two time measurement points for a subject.

In terms of computer use, it was found that all subjects use their computers daily or at least several times a week. In the case of Internet use, it was found that all subjects use the Internet on a daily basis.

A test for normal distribution of the OLLES dimensions revealed that the dimensions Student Collaboration (SC), Information Design and Appeal (IDA), Material Environment (ME), and Reflective Thinking (RT) are normally distributed and the dimensions Computer Competence (CC), Active Learning (AL), and Tutor Support (TS) are not normally distributed. Those descriptive values can be seen in Table 1.

A test for normal distribution for the dimensions of the IPQ revealed that the General Presence (G), Spatial Presence (SP), and Involvement (INV) variables were normally distributed, and the Experienced Realism (REAL) variable was not normally distributed. Those descriptive values can be seen in Table 2.

Next are the results of the qualitative questionnaire. A complete overview of the guideline interview can be found in our previous paper [1] and can be referred to for better understanding. Question 1 revealed that all subjects could still remember the seminar and the use of gather.town well to very well. Question 2 revealed that cooperation within gather.town was rated as sufficient to good. Walking around and interaction opportunities were rated positively. Beyond that, however, additional tools for collaborative workshops like, e.g., Miro [19] outside from the gather.town environment were more likely to be used. Nevertheless, further inquiry revealed that most subjects indicated that there was enough opportunity for successful collaboration. However, some also said that it was somewhat difficult for them to assess this, since they had only used a few functions themselves. Question 3 showed that although there were sometimes technical problems in using gather.town, as an example the browser compatibility, the use itself was always understandable and simple and therefore it did not represent a technical hurdle. Question 4 showed mixed responses. Some subjects found gather.town motivating because it has a certain gaming character and thus offers more functions and possibilities than Zoom, for example. On the other hand, however, it was also increasingly noted that concentration suffers in online seminars and a general demotivation takes place, since the exchange is missing and the classroom is generally preferred. This was also confirmed by the query. Walking around independently in gather.town is more motivating than Classic Video Conferencing Tools, but more demotivating than a seminar in a real classroom. Question 5 and the related query revealed that the tutor's contact and accessibility was good and enough opportunities were given for feedback, and further questions were answered quickly. Based on question 6, it was found that the learning materials were perceived in a very mixed way. However, the query showed that the learning environment apart from the learning materials was perceived as very interesting and appealing. Especially the "Pokémon charm" was very appealing and cute. Question 7 showed that the test persons assess their learning success minimally better

 ISBN: 978-1-68558-089-6

TABLE I. OLLES – STRATEGY IN GATHER.TOWN

Descriptive Analysis					
Dimension	***Mean Value***	***Standard Error of the Mean***	***Standard Deviation***	***Minimum Value***	***Maximum Value***
Student Collaboartion (SC)	3.33	0.22	0.87	1.60	4.60
Computer Competence (CC)	4.71	0.12	0.48	3.50	5.00
Active Learning (AL)	3.39	0.15	0.58	2.60	4.50
Tutor Support (TS)	3.92	0.11	0.43	3.40	5.00
Information Design and Appeal (IDA)	3.49	0.13	0.52	2.70	4.50
Material Environment (ME)	4.00	0.14	0.57	3.00	5.00
Reflective Thinking (RT)	3.13	0.23	0.57	3.00	5.00

TABLE II. IPQ – STRATEGY IN GATHER.TOWN

Descriptive Analysis					
Dimension	***Mean Value***	***Standard Error of the Mean***	***Standard Deviation***	***Minimum Value***	***Maximum Value***
General Presence (G)	1.69	0.34	1.38	0.00	4.50
Spatial Presence (SP)	2.69	0.18	0.70	1.40	3.60
Involvement (INV)	2.04	0.11	0.43	1.38	2.88
Experienced Realism (REAL)	1.78	0.11	0.45	1.25	2.75

than with Classic Video Conferencing Tools, however, they generally assess their learning success online lower than in presence, even if the test persons think that this does not necessarily have an effect on the grades, nevertheless felt on the knowledge that remains at the end. Question 8 further confirmed that subjects prefer gather.town over Classic Video Conferencing Tools like Zoom because it offers more interaction options, facilitates individual conversations, it is very easy to log in, and is generally more dynamic. However, there was also one respondent who preferred Zoom simply out of habit. Question 9 then went on to confirm that all subjects preferred face-to-face lectures. The main reasons for this are that one can interact best with each other, there is also a physical exchange with people, it is more personal and one is less distracted than at home. Question 10 showed that some subjects also used gather.town outside of the lecture for quick communication for projects, or in the work context. However, some did not continue to use it. Finally, question 11 and the two follow-up questions showed that it would be better to integrate additional tools, but gather.town was generally well received due to the diversity as well as physical activation (e.g. yoga) and provides a lot of potential for creative things. Isolated connection problems and browser incompatibility were mentioned as negative points.

B. Results for "Trend" using Zoom

First, the Wilcoxon test will be used to examine whether there are differences in the OLLES test between the individual time measurement points and thus whether there was a change in the evaluation with regard to the repetition of the use of Zoom.

Two time measurement points were not available for all 17 subjects, therefore, the following Wilcoxon test was only calculated with n = 10 complete subjects.

The Wilcoxon test showed that there was no difference between time measurement point 1 and time measurement point 2 regarding the OLLES questionnaire.

Next, using the Wilcoxon test will be used to examine whether there are differences in the IPQ test between the individual time measurement points and thus whether there was a change in the evaluation with regard to the repetition of the use of the Zoom.

The Wilcoxon test showed that there was no difference between time measurement point 1 and time measurement point 2 regarding the IPQ.

The next step is a descriptive analysis of the mean value variables of both time measurement points together. This also includes all measured values regardless of whether there were only one or two time measurement points for a subject.

In terms of computer use, it was found that all subjects use their computers daily or at least several times a week. In the case of Internet use, it was found that all subjects use the Internet on a daily basis.

A test for normal distribution of the OLLES dimensions revealed that the dimensions Student Collaboration (SC), Information Design and Appeal (IDA), Material Environment

 ISBN: 978-1-68558-089-6

(ME), and Reflective Thinking (RT) are normally distributed and the dimensions Computer Competence (CC), Active Learning (AL), and Tutor Support (TS) are not normally distributed. Those descriptive values can be seen in Table 3.

A test for normal distribution for the dimensions of the IPQ revealed that the General Presence (G), Spatial Presence (SP), and Involvement (INV) variables were normally distributed, and the Experienced Realism (REAL) variable was not normally distributed. Those descriptive values can be seen in Table 4.

Next are the results of the qualitative questionnaire. A complete overview of the guideline interview can be found in our previous paper [1] and can be referred to for better understanding. Question 1 revealed that all subjects could still remember the seminar and the use of Zoom well to very well. Question 2 revealed that cooperation within Zoom was rated as sufficient to good. Further inquiry revealed that most subjects indicated that there was enough opportunity for successful collaboration. However, there were problems with collaboration due to a lack of a personal level, which was especially exacerbated by cameras being turned off. Question 3 showed that there were no technical problems in using Zoom and the use itself was always understandable and simple. Question 4 revealed that the use of Zoom mostly demotivated the subjects. One respondent, however, stated that he was more motivated because of the time saved. Time saving was more often mentioned as a positive point while less involvement and more distraction at home were mentioned as negative points. One respondent therefore also felt that the sense of learning together is lost somewhere. Question 5 and the related query revealed that the tutor's contact and accessibility was good and enough opportunities were given for feedback, and further questions were answered quickly. Based on question 6, it was found that the learning materials were perceived in a very mixed way. However, the query showed that the learning environment apart from the learning materials was perceived as very neutral, sometimes even boring, but sufficient to fulfill the purpose. Question 7 showed that the test persons assess their learning success much worse than in presence. Only one respondent stated that he might have even better learning success than in presence, because this allowed him to focus exclusively on the learning content. Based on question 8, a mixed opinion emerged. Some subjects prefer Zoom because Zoom contains fewer distractions from game-like elements. Exactly the opposite, some subjects prefer gather.town because of playful elements, as these promote interpersonal relationships and group work. It was often said that the more interactive and intensive the group work, the more likely they would choose gather.town, but Zoom is perfectly adequate for normal lectures. Question 9 then went on to confirm that most subjects preferred face-to-face lectures. The main reasons for this are that it is more

TABLE III. OLLES - TREND IN ZOOM

Descriptive Analysis					
Dimension	***Mean Value***	***Standard Error of the Mean***	***Standard Deviation***	***Minimum Value***	***Maximum Value***
Student Collaboartion (SC)	3.29	0.20	0.82	1.00	4.40
Computer Competence (CC)	4.69	0.13	0.54	3.00	5.00
Active Learning (AL)	2.96	0.15	0.63	2.00	4.00
Tutor Support (TS)	3.86	0.11	0.43	3.00	4.60
Information Design and Appeal (IDA)	3.10	0.15	0.64	1.60	3.80
Material Environment (ME)	3.75	0.17	0.72	1.40	4.50
Reflective Thinking (RT)	3.18	0.23	0.96	1.20	4.90

TABLE IV. IPQ - TREND IN ZOOM

Descriptive Analysis					
Dimension	***Mean Value***	***Standard Error of the Mean***	***Standard Deviation***	***Minimum Value***	***Maximum Value***
General Presence (G)	0.82	0.32	1.31	0.00	4.00
Spatial Presence (SP)	2.45	0.19	0.78	1.20	4.00
Involvement (INV)	1.87	0.12	0.49	1.50	3.50
Experienced Realism (REAL)	1.86	0.12	0.49	1.38	3.00

 ISBN: 978-1-68558-089-6

personal, and they prefer the physical exchange with people before and after a lecture. They can also pay more attention when they are present, and they are less likely to be distracted. However, one respondent also prefers Zoom because of the time and cost savings in particular. Question 10 showed that some subjects also used Zoom outside of the lecture for projects, or in the work context. Others used it only during the time of the COVID-19 pandemic. Finally, question 11 and the two follow-up questions showed that Zoom is simple, runs stably and the important functions are well integrated. However, it is easier to sit back and turn off the cameras, and this means that the group loses a lot.

C. Group Comparisons

The following is a comparison of all three seminars conducted to date. These are the seminar Trend held in gather.town [1], the seminar Strategy held in gather.town and the seminar Trend held in Zoom. For an overview, see Figure 1. First, the results of the OLLES and the IPQ questionnaire are compared. Afterwards, the exam grades are compared as a performance measure.

First, the seminar Trend (gather.town) was tested with the seminar Strategy (gather.town). In total, the data of 32 subjects are compared. There are 16 from Trend (gather.town) and 16 from Strategy (gather.town). The Mann-Whitney U test showed no significance for any variable. The computer and Internet variables remained without significant difference, as did the OLLES variables. The IPQ could not be tested here, because no survey of the IPQ was conducted for the seminar Trend (gather.town).

Second, the seminar Trend (gather.town) was tested with the seminar Trend (Zoom). In total, 33 subjects are compared. There are 16 from Trend (gather.town) and 17 from Trend (Zoom). Here, the IPQ also could not be tested. The Mann-Whitney U test showed a significant difference in the Active Learning (AL) and Information Design and Appeal (IDA) variables of the OLLES.

Subjects in the gather.town learning environment perceive Active Learning (Mdn = 3.6) better than subjects in the Zoom learning environment (Mdn = 3.0), asymptotic Mann-Whitney U test: U = 57.000, p = .004. Cohen's effect size is r = .50, corresponding to a strong effect.

Subjects in the gather.town learning environment perceive the Information Design and Appeal (Mdn = 3.6) better than subjects in the Zoom learning environment (Mdn = 3.2), asymptotic Mann-Whitney U test: U = 57.000, p = .004. Cohen's effect size is r = .50, corresponding to a strong effect.

Last, the seminar Strategy (gather.town) was tested with the seminar Trend (Zoom). In total 33 subjects are compared. There are 16 from Strategy (gather.town) and 17 from Trend (Zoom). The Mann-Whitney U test showed only a significant difference in the General Presence (G) variable of the IPQ.

Subjects in the gather.town learning environment perceived General Presence (Mdn = 1.25) better than subjects in the Zoom learning environment (Mdn = .00), asymptotic Mann-Whitney U test: U = 73.000, p = .019. Cohen's effect size is r = .41, corresponding to a medium effect.

When comparing grades, the seminar Trend (gather.town) is compared with the seminar Trend (Zoom) first. In total 36 subjects are compared. There are 17 from Trend (gather.town) and 19 from Trend (Zoom). The Mann-Whitney U test showed a significant difference.

Subjects in the gather.town learning environment have better grades (Mdn = 1.7, low values represent better grades) than subjects in the Zoom learning environment (Mdn = 1.9), asymptotic Mann-Whitney U test: U = 90.000, p = .021. Cohen's effect size is r = .38, corresponding to a medium effect.

IV. Discussion

In the dimensions of computer use and Internet use, the subjects indicated that they use this on a daily basis. In addition, the gather.town environment as well as the Zoom environment and all basic functions were sufficiently explained before the start of the study. Thus, we assume that there were no poor ratings for the environments due to possible lack of technical skills.

The test whether there were differences between different time measurement points showed the following results. With the Strategy seminar and the OLLES questionnaire there was no difference in the time measurement points and with the IPQ, there was a difference in scale G (General Presence). The difference in scale G could be explained by the fact that it consists of only one question item and therefore reacts much more strongly to minimal deviations. At the seminar Trend, no significant difference was found between the two time measurement points for either the OLLES or the IPQ. Although a meta-study by Merchant et al. [20] found small effects in simulation studies in terms of number of sessions, these were measures of learning outcome and not an assessment of the immersive environment as in this study. In our previous paper [1], there were also no significant differences at several different time measurement points. Therefore, it can be assumed that it is sufficient to query the questionnaires once.

If one compares the statements of the qualitative questionnaires, it becomes clear that the same statements can be found repeatedly. Almost all subjects showed a hierarchy in their preferred choice of teaching styles. Classroom teaching is clearly preferred. This is followed by the use of 2D Virtual Environments. Classic Video Conferencing Systems are least preferred. If we take a closer look at this hierarchy, we can see that the more opportunities for interaction and the more personal a teaching style is, the more it is preferred. Subjects consistently said they preferred gather.town over Zoom because they had more human proximity and also more opportunities to interact with other students. Nevertheless, ideally, they would like face-to-face teaching. This statement seems to be even more prevalent after the COVID-19 pandemic. However, it also became clear that simple lectures could be replaced more easily by online teaching than seminars in which the focus is on working together.

The group comparisons showed that a comparison of two different seminars with different subjects in gather.town nevertheless resulted in equal evaluations of the Virtual Learning Environment regarding the OLLES questionnaire. Therefore, stable valuations can be assumed here. A comparison of the same seminar with different Virtual

 ISBN: 978-1-68558-089-6

Learning Environments showed that gather.town scored significantly higher on the Active Learning (AL) and Information Design and Appeal (IDA) dimensions of the OLLES questionnaire than Zoom. However, this result could not be repeated for different seminars and different Virtual Learning Environments. There was a significant difference in the G scale of the IPQ, with gather.town showing a higher general presence than Zoom. The Active Learning (AL) dimension of the OLLES specifically asks about the motivation created, as well as the feedback received through the activities or the teaching unit within the environment itself. Again, various studies already showed that motivation [21] - [25] is a crucial factor in the use of VLE's. That there was increased motivation was confirmed by the interviews. The motivation arose primarily through increased interactivity. For the test persons, it was clearly more motivating to walk through the Virtual Environment by moving the avatar and not just to sit in front of the laptop. This also led to the environment being perceived as very varied. The dimension Information Design and Appeal (IDA) of the OLLES asks in particular how creative and original presented teaching materials are and whether graphics used are helpful and visually appealing. This mainly refers to the teaching slides presented as if they were in a presentation. Since the same learning materials were used here, this difference is difficult to explain. It is possible that the actual learning environment was included in the evaluation and not just the learning materials. Perhaps this double assessment was due to the fact that, in this particular case, it was not always clear to the subjects what the individual question items referred to in this dimension. The scale G (General Presence) of the IPQ asks solely about the sense of being there. This feeling could not be created at all with Zoom and at least minimally with gather.town. However, only in one of the two tests with different seminars. Whether there is an influence of the seminar on the evaluation of a Virtual Learning Environment is difficult to say. Nevertheless, the results found could also be due to a still small sample size. Statistically, however, the difference between the two Virtual Learning Environments turned out to be smaller than the qualitative interviews suggested. In the end, only partially significant differences in the evaluation could be found and these could not be repeated.

Looking at the exam grades, a significant difference was found between the Virtual Learning Environments used. When using the gather.town environment, the subjects had better grades than using the Zoom environment. This is a medium effect. Although there was not much difference in the assessment of Virtual Learning Environments, it does seem to have an impact on performance measurement in the form of exam grades. The results also confirmed that it is only possible to compare the same seminars with each other.

V. Conclusion And Future Work

This study shows that, according to the subjects, there is a hierarchy of teaching styles. Classroom teaching is the most popular form. This is due to the direct contact with fellow students, greater motivation and the best possible opportunity for interaction in order to solve tasks in a team and learn together. This is followed by the use of a 2D Virtual Learning Environment. Here, direct contact is much more limited than in face-to-face teaching, but this can be partially replaced by the use of avatars and the resulting interaction possibilities. Thus, the test participants are also motivated to use the Virtual Learning Environment. The most unpopular are Classic Video Conference Systems. These have the least interaction possibilities and are therefore perceived as demotivating. This hierarchy, especially the preference of face-to-face personal teaching is confirmed by several other studies [26] - [29]. Also the preference for gather.town as 2D Desktop VR to Zoom as Classic Video Conferencing Tool can be explained and

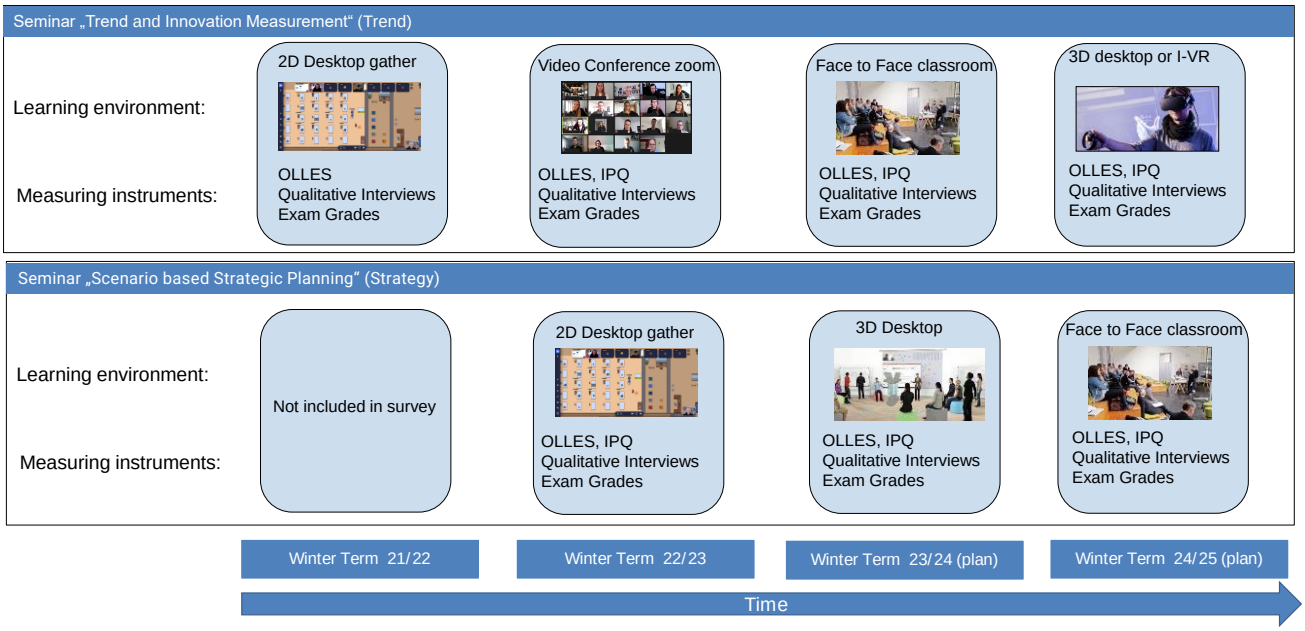

Figure 4. Overview of seminars, learning environments and measuring instruments for finished and planned studies.

confirmed by several studies [30] - [33]. It seems to be important to use VLE that are innovative, social emotional, and engage formal and informal communication, which seems to be better solved within the Virtual 2D Learning Environment gather.town.

In an evaluation of Virtual 2D Learning Environments and Classic Video Conference Systems using the OLLES and IPQ questionnaire, however, this could only be shown for some dimensions or scales. Contrary to the statements of the qualitative interviews, the quantitative evaluation of the two online teaching formats therefore seems to make no or only a very small difference. In contrast, when exam grades were measured as a performance measure, subjects were found to perform better with Virtual 2D Learning Environments than with Classic Video Conference Systems. Thus, the use of 2D Virtual Learning Environments seems to be a better choice than Classic Video Conference Systems for successful online teaching. However, it must also be noted here, that this is a field study and therefore the number of subjects is low. Future work needs to clarify whether face-to-face teaching also leads to the best performance measures. In addition, other online forms of teaching will also be tested. For this purpose, it is initially planned to hold the same seminars as in this study in the next semester once in face-to-face teaching and once in a Virtual 3D Learning Environment. An overview can be seen in Figure 4.

Since it has been found that realism plays an important factor in the evaluation of Virtual Learning Environments, this will also be used to explore which factors contribute to a higher degree of realism. For example, the change from a 2D learning environment to a 3D learning environment with 3D avatars could be an improvement. This could then be seen with a better IPQ rating. In addition, this study will be extended to the application of I-VR environments, as soon as this can be implemented with enough test persons, since sufficient equipment must be available and software must offer all necessary functions. Now, there are many indications that hybrid forms of teaching and learning will be used in the future. Above all, the type of seminar also plays a role. Roughly speaking, the more interactive the seminar, the more opportunities for interaction are required and the more the seminar should tend towards classroom teaching. It also shows that personal contact cannot be replaced and that this provides more motivation for learning. In the end, the goal should always be to provide the best possible teaching and learning experience for all involved.

REFERENCES

[1] G. Hube, K. Pfeffel, and N. H. Müller, **"**2D Virtual Learning Environments for Tertiary Education" International Journal on Advances in Systems and Measurements, ISSN 1942-261x, vol. 15, no. 3 & 4, pp. 81-92, 2022.

[2] J. Clayton, "Development and Validation of an Instrument for Assessing Online Learning Environments in Tertiary Education: The Online Learning Environment Survey (OLLES)," [Online]. Available from: https://espace.curtin.edu.au/handle/20.500.11937/550 2023.10.10

[3] T. Schubert, F. Friedmann, and H. Regenbrecht, "The experience of presence: Factor analytic insights," Presence, vol. 10, no. 3, pp. 266–281, 2001, doi: 10.1162/105474601300343603.

[4] G. Q. Hu, "Qualitative Analysis of Students' Online Learning Experiences after the University Reopening," Journal of Education, Humanities and Social Sciences, vol. 7, pp. 115–134, Jan. 2023, doi: 10.54097/ehss.v7i.4074.

[5] I. Assaly and U. Atamna, "Who Needs Zoom? Female Arab Students' Perceptions of Face-to-Face Learning and Learning on Zoom," Sustainability, vol. 15, no. 10, 8195, 2023.

[6] N. Kocyigit and F. Yilmaz, "Effects of Zoom Fatigues on Life Satisfaction: A Research on Teachers," [Online]. Available from: https://www.researchgate.net/publication/371970282_EFFECTS_OF_ZOOM_FATIGUE%27S_ON_LIFE_SATISFACTION_A_RESEARCH_ON_TEACHERS 2023.10.10

[7] P. Prasetyo and Z. Abidin, "Zoom Learning Media Relatioship and Interest in Learning with Learning Outcomes Civics," Akademika: Jurnal Teknologi Pendidikan, vol. 12, no. 1, pp. 153-161, 2023. doi: 10.34005/akademika.v12i01.2467

[8] L. Andrade-Arenas, W. W. Reyes Perez, and C. A. Yactayo Arias, "Moodle platform and Zoom videoconference: learning skills in the virtual modality," Indonesian Journal of Electrical Engineering and Computer Science, vol. 31, no. 1, pp. 337-349, 2023, doi: 10.11591/ijeecs.v31.i1.pp337-349.

[9] A. Carțiș, ""Zoom Fatigue" In Higher Education: Videoconferencing Impact On Students' Fatigue," Education Facing Contemporary World Issues - EDU WORLD 2022, vol. 5, pp. 1355-1364, 2023, doi: 10.15405/epes.23045.138.

[10] L. Knox, S. Berzenski, and S. Drew, "Measuring Zoom Fatigue in College Students: Development and Validation of the Meeting Fatigue Scale for Videoconferencing (MFS-V) and the Meeting Fatigue Scale for In-Person (MFS-I)," Media Psychology, Advance online publication, doi: 10.1080/15213269.2023.2204529

[11] A. Ngien and B. Hogan, "The relationship between Zoom use with the camera on and Zoom fatigue: considering self-monitoring and social interaction anxiety," Information Communication & Society, vol. 26, no. 10, pp. 2052-2070, 2023, doi: 10.1080/1369118X.2022.2065214.

[12] G. Fauville, M. Luo, A. C. M. Queiroz, A. Lee, J. N. Bailenson, and J. Hancock, "Video-conferencing usage dynamics and nonverbal mechanisms exacerbate Zoom Fatigue, particularly for women," Computers in Human Behavior Reports, vol. 10, 2023, doi: 10.1016/j.chbr.2023.100271.

[13] H. N. Shoshan and W. Wehrt, "Understanding Zoom Fatigue: A Mixed-Method Approach," Applied Psychology, vol. 71, no. 3, pp. 827-852, 2022, doi: 10.1111/apps.12360.

[14] C. K. Lo and Y. Song, "A Scoping Review of Empirical Studies in Gather.town," 11th International Conference on Information and Education Technology (ICIET), 2023, pp. 1-5, Electronic ISBN: 978-1-6654-6548-9 doi: 10.1109/ICIET56899.2023.10111430.

[15] Gather Presence, Inc. gather.town. [Online]. Available from: https://www.gather.town 2023.10.10

[16] D. Serhan, "Transitioning from face-to-face to remote learning: Students' attitudes and perceptions of using Zoom during COVID-19 pandemic," International Journal of Technology in Education and Science, vol. 4, no. 4, pp. 335-342, 2020, doi: 10.46328/ijtes.v4i4.148.

[17] Zoom Video Communications, Inc. [Online]. Available from: www. https://zoom.us 2023.10.10

[18] R. Likert, "A technique for the measurement of attitudes," Archives of Psychology, vol. 22, no. 140, pp. 5-55, 1932.

[19] Miro. [Online]. Available from: https://miro.com,f14 2023.07.11

[20] Z. Merchant, E. T. Goetz, L. Cifuentes, W. Keeney-Kennicutt, and T. J. Davis, "Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis," Computers & Education, vol. 70, no. 1, pp. 29–40, 2014, doi: 10.1016/j.compedu.2013.07.033.

[21] S. Y. Chien, G. J. Hwang, and M. S. Y. Jong, "Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions," Computers & Education, vol. 146, no. 3, 2019, doi: 10.1016/j.compedu.2019.103751.

[22] M. H. Kim, "Effects of Collaborative Learning in a Virtual Environment on Students' Academic Achievement and Satisfaction," Journal of Digital Convergence, vol. 19, no. 4, pp. 1–8, 2021, doi: 10.14400/JDC.2021.19.4.001.

[23] B. Yildirim, E. Sahin-Topalcengiz, G. Arikan, and S. Timur, "Using virtual reality in the classroom: Reflections of STEM teachers on the use of teaching and learning tools," Journal of Education in Science, Environment and Health, vol. 6, no. 3, pp. 231-245, 2020, doi: 10.21891/jeseh.711779.

[24] S. Mystakidis, E. Berki, J. P. Valtanen, "Deep and Meaningful E-Learning with Social virtual reality Environments in Higher Education: A Systematic Literature Review," Applied Sciences, vol. 11, no. 5, 2412, 2021, doi: 10.3390/app11052412.

[25] M. Akgün, B. Atıcı, "The Effects of Immersive virtual reality Environments on Students' Academic Achievement: A Meta-analytical and Meta-thematic Study," Participatory Educational Research, vol. 9, no. 3, pp. 111-131, 2022. doi: 10.17275/per.22.57.9.3.

[26] A. Driscoll, K. Jicha, A. N. Hunt, L. Tichavsky, and G. Thompson, "Can online courses deliver in-class results?: A comparison of student performance and satisfaction in an online versus a faceto-face introductory sociology course," Teaching Sociology, vol. 40, no. 4, pp. 312–331, 2012, doi: 10.1177/0092055X12446624.

[27] C. Stone, "Online learning in Australian higher education: Opportunities, challenges and transformations," Student Success, vol. 10, no. 2, pp. 1–11, 2019, doi: 10.5204/ssj.v10i2.1299.

[28] C. Merlin-Knoblich, P. N. Harris, and E. C. McCarty Mason, "Examining student classroom engagement in fipped and non-fipped counsellor education courses," The Professional Counselor, vol. 9, no. 2, pp. 109–125, 2019, doi: 10.15241/cmk.9.2.109.

[29] M. R. Cairns, M. Ebinger, C. Stinson, and J. Jordan, "COVID-19 and human connection: Collaborative research on loneliness and online worlds from a socially-distanced academy," Human Organization, vol. 79, no. 4, pp. 281–291, 2020, doi: 10.17730/1938-3525-79.4.281.

[30] J. Du, X. Fan, J. Xu, C. Wang, L. Sun, and F. Liu, "Predictors for students' self-efcacy in online collaborative groupwork," Educational Technology Research and Development, vol. 67, pp. 767–791, 2019, doi: 10.1007/s11423-018-9631-9.

[31] L. Caprara and C. Caprara, "Effects of virtual learning environments: A scoping review of literature," Education and Information Technologies, vol. 27, pp. 3683–3722, 2022, doi: 10.1007/s10639-021-10768-w.

[32] S. V. Mamadjanova, "Design Features of Virtual Learning Environments," European International Journal of Multidisciplinary Research and Management Studies, vol. 2, no. 6, pp. 1–5, 2022, doi: 10.55640/eijmrms-02-06-01.

[33] C. Latulipe and A. De Jaeger, "Comparing Student Experiences of Collaborative Learning in Synchronous CS1 Classes in Gather.Town vs. Zoom," SIGCSE 2022: Proceedings of the 53rd ACM Technical Symposium on Computer Science Education - vol. 1, Feb. 2022, pp. 411-417, doi: 10.1145/3478431.3499383.

 ISBN: 978-1-68558-089-6