ASE

Immersive Anatomy Atlas: Learning Factual Medical Knowledge in a Virtual Reality Environment

Kilian Gloy,[1]* Paul Weyhe,[1] Eric Nerenz,[1] Maximilian Kaluschke,[2] Verena Uslar,[1] Gabriel Zachmann,[2] Dirk Weyhe[1]

[1]Department for Visceral Surgery, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany
[2]Institute for Computer Graphics and Virtual Reality, University of Bremen, Bremen, Germany

In order to improve learning efficiency and memory retention in medical teaching, furthering active learning seems to be an effective alternative to classical teaching. One option to make active exploration of the subject matter possible is the use of virtual reality (VR) technology. The authors developed an immersive anatomy atlas which allows users to explore human anatomical structures interactively through virtual dissection. Thirty-two senior-class students from two German high schools with no prior formal medical training were separated into two groups and tasked with answering an anatomical questionnaire. One group used traditional anatomical textbooks and the other used the immersive virtual reality atlas. The time needed to answer the questions was measured. Several weeks later, the participants answered a similar questionnaire with different anatomical questions in order to test memory retention. The VR group took significantly less time to answer the questionnaire, and participants from the VR group had significantly better results over both tests. Based on the results of this study, VR learning seems to be more efficient and to have better long-term effects for the study of anatomy. The reason for that could lie in the VR environment's high immersion, and the possibility to freely and interactively explore a realistic representation of human anatomy. Immersive VR technology offers many possibilities for medical teaching and training, especially as a support for cadaver dissection courses. Anat Sci Educ 15: 360–368. © 2021 The Authors. Anatomical Sciences Education published by Wiley Periodicals LLC on behalf of American Association for Anatomy.

Key words: gross anatomy education; anatomical atlas; virtual reality; immersive VR; active learning; knowledge retention

INTRODUCTION

Medical education not only requires the trainees to acquire a number of practical skills, but also to learn large amounts of basic factual information. This circumstance makes efficient learning and accurate retention imperative (Yeh and Park, 2015). Many different methods have been explored in order to improve learning efficiency (Yeh and Park, 2015). One type of learning that is consistently connected with improved learning and memory results is active learning (Hazlett, 2009; Kornell et al., 2009; Markant et al., 2016). There are several different definitions of active learning, from different fields of study; Common aspects of those definitions include: "some combination of increased physical activity or interaction, deeper processing, elaboration or explanation of material, planning of learning activities, question asking, metacognitive monitoring, and social collaboration" (Markant et al., 2016). These criteria tie well into the psychological theory of constructivist learning which assumes learning to be an active process and learners to be actively seeking knowledge. In complex interaction with the material, the learners generate knowledge and deeper understanding (Fosnot and Perry, 1996; Siemens, 2005).

In recent years, the use of virtual reality (VR) technology has been considered as a useful tool in education (Markant et al., 2016). In line with the ideas of active and constructivist

*Correspondence to: Dr. Kilian Gloy, Abteilung für Viszeralchirurgie, Carl von Ossietzky University of Oldenburg, Grüne Str. 6, 26121 Oldenburg, Germany. E-mail: kilian.gloy@uni-oldenburg.de

Additional supporting information can be viewed in the online version of this article.

Received 13 July 2020; Revised 16 April 2021; Accepted 19 April 2021.

Published online 30 June 2021 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/ase.2095

learning, VR could be used to increase the physical interaction with the subject matter and allow learning in an explorative context more similar to real-life conditions. The additional control, which learners would have over the experience, can be expected to improve learning (Gureckis and Markant, 2012), and the digital environment could also allow for gamification in order to increase learner motivation (Koivisto and Hamari, 2019). It has already been shown that VR has a positive impact on learning compared to conveying information via desktop personal computer (Selzer et al., 2019), likely due to the increased immersion which aids in information recall (Krokos et al., 2019).

The immersion VR provides is highest for the so-called "immersive VR", in which a user can interact with a computer-generated three-dimensional (3D) environment as if they were physically present in that environment (Freina and Ott, 2015; Zackoff et al., 2019). Immersive VR is associated with higher ratings for interest and motivation in students (Parong and Mayer, 2018). Despite the advantages of immersive VR, there are two potentially negative aspects: cybersickness and high cognitive load.

Cybersickness is a phenomenon common in interaction with virtual environments, especially VR, and consists of a multitude of physiological symptoms similar to car- or seasickness (Brewer-Deluce et al., 2021). The cause of this unpleasant experience is unclear, but likely related to sensory mismatch (Yildirim, 2019). However, studies which used immersive VR in surgical training have reported little to no problems with cybersickness in their participants (Huber et al., 2017; Frederiksen et al., 2020). Frederiksen and colleagues argued that this may be because of the limited head movements in this setting compared to average VR games (Frederiksen et al., 2020).

The cognitive load of a learning task is commonly divided into at least two types: the intrinsic load, inherent to the task or information that must be learned, and the extraneous load, generated by external processes or information that distract from the learning material (Wong et al., 2012). There is evidence that, in laparoscopic surgical training, immersive VR has increased cognitive load and an associated worsened task performance (Frederiksen et al., 2020). The complete implications of these circumstances have yet to be discussed; For example, it was argued that training under increased cognitive load may actually be beneficial since it improves the transfer of training into a real situation with a strong cognitive load (Sankaranarayanan et al., 2020).

Within the medical field, VR technology has been shown to be a helpful tool for teaching procedural skills (Bracq et al., 2019), since the proficiency acquired there can transfer to the real-world clinical setting (Seymour, 2008). So far, it seems that conventional VR outperforms immersive VR in this context (Frederiksen et al., 2020). Medical trainees have to acquire both procedural skills and factual knowledge. There have been a few research studies into the effectiveness of immersive VR for the teaching of anatomy, but a clear result has yet to emerge. For example, Stepan and colleagues found that VR provided a more enjoyable learning experience than textbooks without actually increasing the learning benefit (Stepan et al., 2017), whereas Kurul and colleagues found a significantly positive learning effect of VR compared to attending a presentation on the material (Kurul et al., 2020). A study by Birbara and colleagues found that the learning preferences differed between different groups of participants (tutors vs. students), although immersive VR was seen as more mentally taxing than a desktop version of the same program, and more strongly connected to physical discomfort (Birbara et al., 2020). Lastly, Zinchenko and colleagues found immersive VR to be most beneficial for learning previously unknown information when compared to books and a 3D desktop application (Zinchenko et al., 2020).

In summary, the current state of research on immersive VR as a tool for learning human anatomy is ambiguous. Further research in different populations and with different methods is necessary to gather more empirical data and piece together the whole picture. Seeing as anatomy study through cadaver dissection has many advantages which neither textbooks nor VR applications can recreate (Dua et al., 2021), it is unlikely that VR will replace this traditional learning method. However, if immersive VR is shown to be effective in learning and retaining anatomy knowledge, it might become a meaningful support in anatomy courses.

The immersive, interactive 3D anatomy atlas used in this study was developed at the VR laboratory at the University of Bremen. The atlas features a virtual operating theater and allows the user to actively explore anatomical structures and arrangements of the human body through virtual dissection. A previous pilot study using an older version of the same atlas has already shown that information acquisition was faster when novices to the study of anatomy used the VR atlas compared to retrieving the information from books (Weyhe et al., 2018). As a second step, the aim of this study was the examination of long-term knowledge retention in novices by measuring the amount of correct answers they give after working with the immersive anatomy atlas (VR condition) in comparison to using only anatomical books (open book condition; OB) utilizing a randomized study design. The ratio of correct answers to a second questionnaire conducted several weeks later operationalized the information retention rate in both groups.

In line with the literature presented above and the aim of this study, the hypotheses were as follows: (1) Acquiring new information with the VR atlas is faster than using standard printed anatomical atlases [replication of effect from (Weyhe et al., 2018)]; and (2) Working with the VR atlas leads to an improved retention of knowledge compared to working with standard anatomical atlases.

MATERIALS AND METHODS

This study was approved by the medical ethics committee of the Carl von Ossietzky Universität Oldenburg, Germany (ID-number: 2020-065).

Immersive Anatomy Atlas

The immersive anatomy atlas is an application developed by the authors that uses a head-mounted display (HMD) to immerse the user in a virtual operation room (see Fig. 1). Using head tracking, users have a full 360 view, that is, they can look around and move within the virtual room. Additionally, there are bi-manual controllers which enable the user to interact with virtual reality. They can manipulate individual organs by grabbing them with their virtual hands, which are controlled by the controllers. Several virtual tools are placed on a nearby table. Some of these tools mimic realistic surgery tools, some allow for more "magic" tasks, such as exploring the anatomy by controlling the model's transparency. Others allow to hide anatomy in spherical areas around a pointing tool (see Fig. 1) or place a cross section to hide all organs in front of it (see Fig. 1).

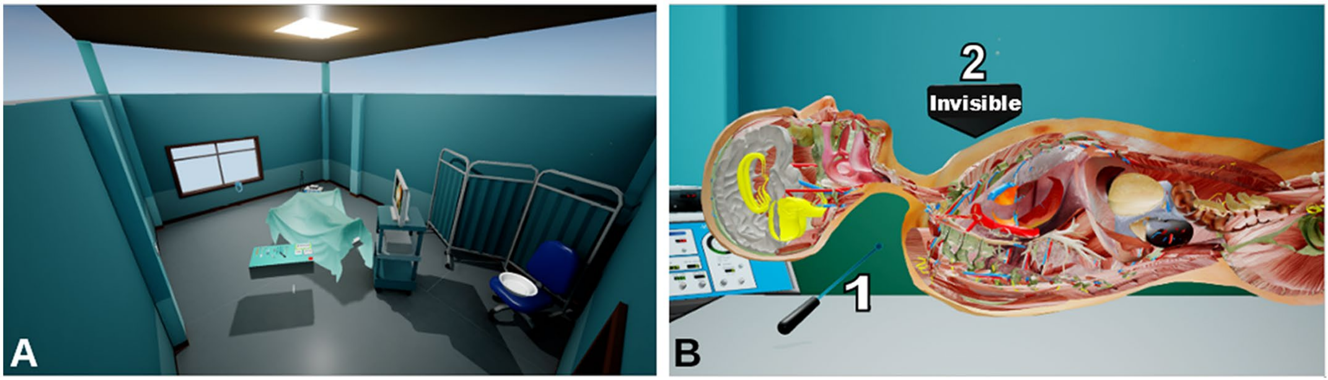

Figure 1.

Immersive Anatomy Atlas. A, Overview of the 3D environment that comprises the virtual operating room. The camera is positioned in one of the rooms upper corners; B, A cross-section of the anatomical model, generated with the use of virtual tools. The pointing baton (1) makes a spherical area around its tip transparent. The cross-sectional tool (2) makes everything on the front side of the imaginary plane which passes through the tool transparent. The combination of these tools enables the user to inspect any part of the model.

During the learning phase, participants can explore and study the immersive anatomy atlas by inspecting it from every point of view, by grabbing organs and other structures and inspecting them from all angles, then replacing them in the original position. When organs are placed back, they snap to their original place, as long as the release pose (position and orientation) is close to their original, correct pose. Thus, the anatomical model is always correct, unless deliberately altered by the user.

While anatomical structures are held in hand, further information about them can be viewed by the user. Furthermore, the complete anatomical model can be reset at once to its original state using a virtual button.

An introductory video for the anatomy atlas used in this study can be viewed online (Pius-Hospital Oldenburg, 2018).

The immersive anatomy atlas system was implemented by the authors on top of the game engine Unreal Engine, version 4.23 (Epic Games, Inc. Cary, NC) using the built-in programming language Blueprint. The 3D geometrical models were created by a 3D artist and purchased by the authors. All the anatomical parts are designed to closely resemble real anatomy. The geometry was further modified by the authors through Blender software, version 2.92.0 (Blender Foundation, Amsterdam, The Netherlands). Coordinate origins were moved, geometry groups were separated, textures were changed, and missing organs were added. Everything is rendered from 3D geometrical models of the anatomical structures at runtime in real time.

The models are loaded at runtime by the immersive anatomy atlas, then rendered stereoscopically by the game engine and displayed in stereo on the head-mounted display (HMD), thus providing stereoscopically correct images to the user. During the study, a head-mounted display HTC VIVE™ (High Tech Computer Corp., New Taipei City, Taiwan) with the resolution of 1080 by 1200 pixel per eye was used to display the immersive anatomy atlas to the users in stereo vision. The frame rate of the immersive anatomy atlas was sustained at 90 frames per second, to allow for the illusion of presence in the virtual reality.

All the virtual tools (both surgery and "magic" ones) were implemented by the authors.

Study Design

The study was conducted at two separate German high schools, referred to as school A and school B below. In each school, 16 participants were recruited and randomly assigned into equally sized groups for two different learning modalities: open book (OB condition) and virtual reality (VR condition) learning (per school: OB: n = 8, VR: n = 8). High school students were chosen to ensure that the participants would have no prior formal anatomical training and would approach the learning content as novices.

A schematic overview of the experimental design can be found in Figure 2.

The VR-group used the immersive anatomy atlas. They viewed a short introductory video for the atlas before the experiment started. They had the opportunity to familiarize themselves with the VR environment for a maximum of 5 minutes and clear up any questions regarding the handling of the VR interface.

The OB-group used standard anatomy atlases (Paulsen and Waschke, 2017a,b,c). The participants in both groups were presented with the same set of nine single-choice questions on paper, encompassing the topics topography, cardiovascular system, and nervous system (see Supplemental Material File 1 for a list of the questions). They were tasked to answer the questions correctly and as quickly as possible, using only the respective method of learning at their disposal (OB or VR). The time which elapsed between the question and the participants' answer was recorded (response time), and the percentage of correct questions constituted the test score. Because of this experimental setup, the questions of this test can't be classified according to Bloom's revised taxonomy (Krathwohl, 2002). As soon as all questions had been answered, either correctly or incorrectly, Test 1 was over (see Fig. 3 for an impression of the experimental set up in Test 1).

After seven weeks in school A and four to five weeks in school B, the participants were tested for their long-term memory of the topics they learned during Test 1. Each participant answered a second list of nine multiple-choice questions from the same three topics as before (see also Supplemental Material File 1). Participants had to answer the questions without any help, based only on their memory. The

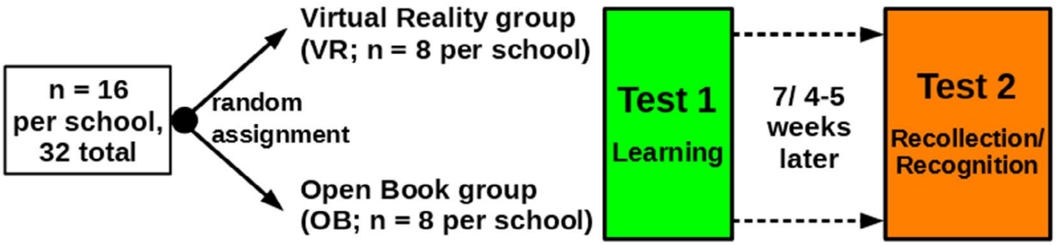

Figure 2.

Schematic overview of the experimental design. Students in both schools were randomly assigned to either VR or OB groups. Afterward, they completed the first test (green) with their respective learning method. Following a seven-week waiting period in school A or a four- to five-week period in school B, the students completed the second test (orange) from memory. VR, virtual reality; OB, open book.

response time was not measured; The sole point of interest was whether the recall of the factual knowledge was successful or not, as reflected in correct or incorrect answers. All questions in this test belong to the category "Remember", according to Bloom's revised taxonomy, since recognizing and recalling information from long-term memory was necessary (Krathwohl, 2002).

The anatomy tests (Test 1 and Test 2) were developed in-house, by medical experts, to ensure that each question can be answered with the given material and is of appropriate difficulty level for high school students. Since the tests were developed specifically for this study and were not tested elsewhere, no statements can be made about the tests' reliability or validity.

Participants

Thirty-two eleventh-grade high school students participated in the experiment on a voluntary basis, 16 from each of the two high schools (overall n = 32). Human anatomy was not part of the senior-classes' biology curriculum, meaning that the courses the participants took during that time had no influence on their preexisting knowledge of anatomy.

In school A, the average age ± SD was 16.5 ± 0.52 years (minimum = 16, maximum = 17,). The OB group consisted of three female and five male students, and the VR group of five female and five male students.

In school B, the average age ± SD was 17.6 ± 0.5 years (minimum = 17, maximum = 18). The OB group consisted of three

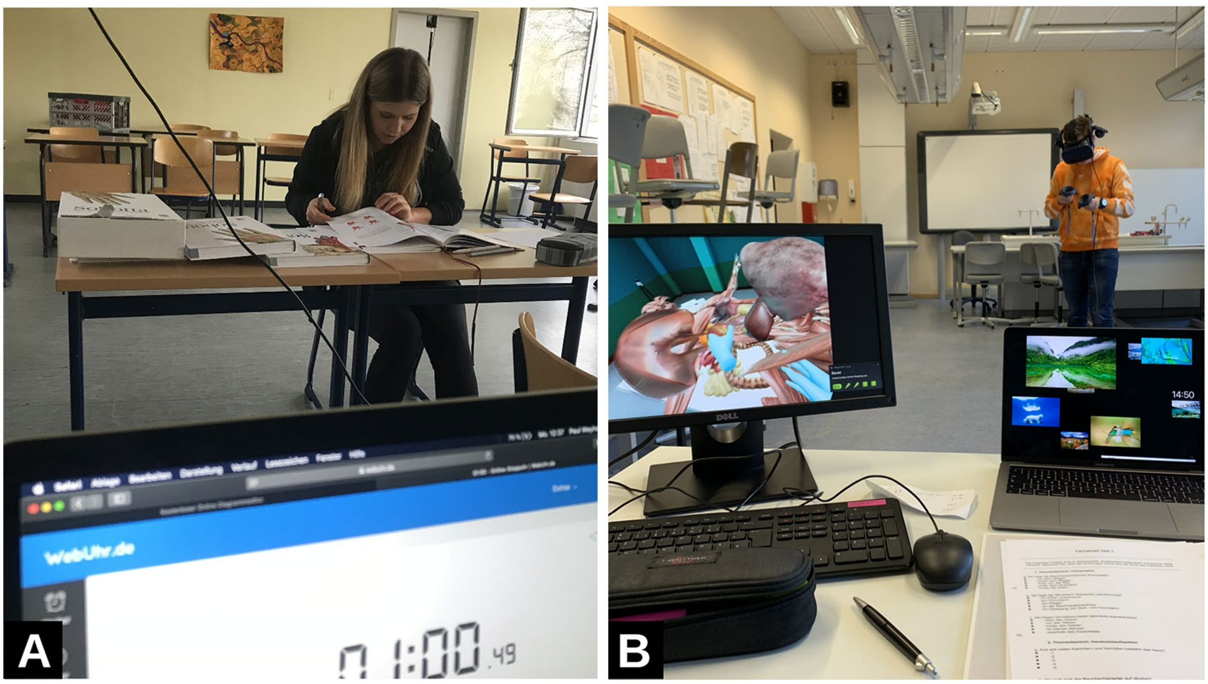

Figure 3.

Impressions of the experimental setup for Test 1. A, The open book condition; A student is using the textbooks supplied to them to answer the questionnaire. A timer can be seen at the lower edge of the panel, measuring response times. B, The virtual reality condition; A student wearing the head-mounted display (HMD) is using the immersive anatomy atlas to answer the questionnaire. A space has been cleared around them to allow for free movement in the virtual reality environment. At the left edge of the panel, the monitor shows what the student is currently seeing in their display.

Table 1.

Results of the *Post Hoc t*-tests for the Mixed ANOVA, for the Interaction Effect of SCHOOL and TIME

		School A		School B	
School	**Test Number**	**Test 1 *P*-value[a]**	**Test 2 *P*-value[a]**	**Test 3 *P*-value[a]**	**Test 4 *P*-value[a]**
School A	Test 1				
	Test 2	<0.001[b]			
School B	Test 3	0.554	<0.001[b]		
	Test 4	<0.001[b]	<0.041[b]	<0.001[b]	

[a]Benjamini–Hochberg corrected *P*-values;[b]Indicates statistically significant results. Empty cells were either nonexistent values or repetitions.

female and five male students, the VR group of four female and four male students.

There were no significant differences in the distribution of gender or age between the learning groups, in neither school, as confirmed with Fischer's exact test and t-test, respectively.

No participant was familiar with the head-mounted display or the immersive anatomy atlas before participating in this study. All participants received a letter detailing the contents of the experiment. Only students who handed in the letter signed by their legal guardians were allowed to participate.

Statistical Analysis

In order to investigate hypothesis 1, a replication of the response time effect found in the preceding study on the immersive anatomy atlas (Weyhe et al., 2018), the response time data from Test 1 in school B were used. School A had to be excluded from this analysis because an error during the experiment led to a loss of the relevant data. The response times were averaged over the questions and then grouped by learning method. The distributions of average response times in both learning groups were tested for normal distribution using the Shapiro–Wilk test ($\alpha = 0.1$), because this test works well for small samples (see Field et al., 2012). Since the assumption of normality was violated, a two-sided Wilcoxon rank sum test had to be used instead of the parametric t-test (Field et al., 2012) to test for an effect of the learning method on the response time.

To test hypothesis 2, the assumption of improved knowledge retention for those working with the immersive anatomy atlas, the test performance was compared between experimental groups, time points, and schools. The performance in the nine individual questions was summarized into one variable that represented the percentage of correctly answered questions per participant and per test. A mixed ANOVA was calculated for the percentage of correct answers with the factors METHOD (VR or OB, between factor), SCHOOL (school A or school B, between factor), and TIME (Test 1 or Test 2, within factor). This statistical test was chosen because three factors had a potential influence on the dependent variable, making an ANOVA necessary in order to make comparisons of means. The assumptions of normality, homoscedasticity, and sphericity were tested beforehand using Shapiro–Wilk tests and Levene tests ($\alpha = 0.1$). The assumption of normality was violated in three of the eight groups, but since the ANOVA is a robust procedure as long as the group sizes are equal (Field et al., 2012) and all other assumptions were fulfilled, the parametric ANOVA was used.

To check the difficulty of the single-choice questions, the difficulty index was calculated for Test 1 and Test 2. Additionally, for an investigation of the individual questions, the performance of all participants was averaged for each of the nine questions, separately for Test 1/Test 2 and VR/OB. The resulting percentage of correct answers per question was then visualized in a bar chart.

All statistical analysis was performed using R statistical software, version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria).

Table 2.

Results of the Post Hoc t-tests for the Mixed ANOVA, for the Interaction Effect of METHOD (Virtual Reality vs. Open Book) and TIME

		Virtual Reality		Open Book	
Method	**Test Number**	**Test 1 *P*-value[a]**	**Test 2 *P*-value[a]**	**Test 3 *P*-value[a]**	**Test 4 *P*-value[a]**
Virtual Reality	Test 1				
	Test 2	<0.001[b]			
Open Book	Test 3	0.175	<0.001[b]		
	Test 4	<0.001[b]	<0.01[b]	<0.001[b]	

[a]Benjamini–Hochberg corrected *P*-values;[b]Indicated significant results. Empty cells were either nonexistent values or repetitions.

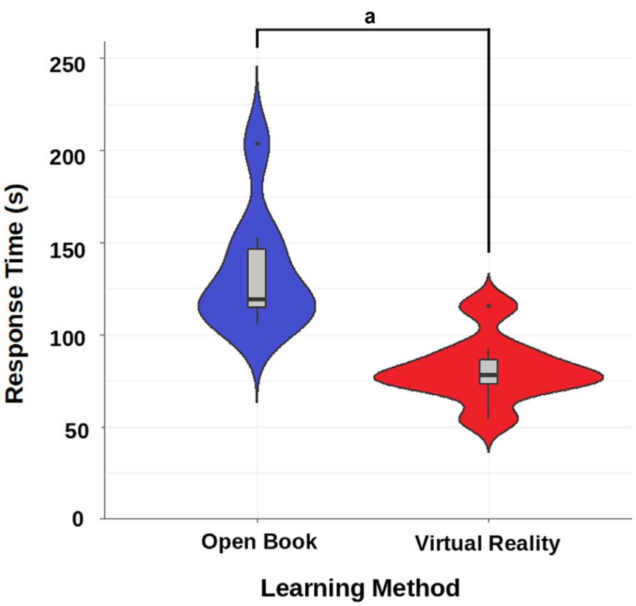

Figure 4.

Response time (in seconds) of Test 1 in school B, grouped by learning method. The colored outlines depict the distribution of the data; therefore, the width of the outline at any point Y increases with the number of participants whose response time results equal Y, or lie close to it. The box plot within shows the quartiles 1, 3, and the median. The black dots denote outliers. [a]The difference between the two methods is significant (P < 0.001).

RESULTS

There was a significant difference between the response times in the OB and the VR learning conditions, as shown by the Wilcoxon rank sum test (median difference = 41.3, $P < 0.001$, $d = 2.01$). It took the participants significantly longer to answer the questions in the OB condition (median = 119.5 seconds) than in the VR condition (median = 78.3 seconds). This difference represents a large effect. The results of the response time analysis are visualized in Figure 4.

There were several significant effects on the performance of the students, operationalized by the percentage of correct answers. The mixed ANOVA showed a main effect of METHOD [$F_{(1, 28)} = 9.15$, $P < 0.01$, partial $\eta^2 = 0.353$]. Participants using the VR atlas achieved better results (mean ± SD of correct answers = 73 ± 22%) than those working with books (mean ± SD of correct answers = 60 ± 25%). The associated effect size represents a large effect.

There was also a main effect of TIME [$F_{(1, 28)} = 113.89$, $P < 0.001$, partial $\eta^2 = 0.803$], and a significant interaction effect SCHOOL and TIME [$F_{(1, 28)} = 6.98$, $P < 0.05$, partial $\eta^2 = 0.199$]. For a visualization of the mixed ANOVA results, see Figure 5.

In order to investigate the interaction effect of SCHOOL and TIME, the mean percentages of correct answers per school were calculated for Test 1 (school A: 85% ± 13%, school B: 82% ± 15%) and Test 2 (school A: 42% ± 16%, school B: 56% ± 21%). Additionally, post hoc t-tests were performed for this interaction effect (Benjamini–Hochberg correction, see Table 1) and, for explorative purposes, the interaction of METHOD and TIME (Benjamini–Hochberg correction, see Table 2).

Test 2 was overall more difficult than Test 1. The difficulty index (in %) for Test 1 was 83.68 (±13.82), with a range of 68.75–100, and for Test 2 it was 48.61 (±19.71), with a range of 9.38–100.

DISCUSSION

The response time effect postulated in hypothesis 1, which was also found in the previous study on the immersive anatomy atlas (Weyhe et al., 2018), was replicated in this study. This was indicated by the significant Wilcoxon rank sum test, in combination with the higher median response time in the OB group compared to the VR group. It follows that acquiring previously unknown information was faster in the VR condition; this confirms hypothesis 1. The reason for this could be the interactive way of retrieving information from the immersive anatomy atlas, which leads to an easier access to factual anatomical knowledge.

The main goal of this study was the investigation of long-term effects of learning through immersive VR. The short-term benefits were already well documented, while little could be said about retention of knowledge over a longer period of time. Now, this study adds the results of the mixed ANOVA on the knowledge-test performance to the relevant empirical evidence. Two main effects of the factors METHOD and TIME were revealed in the ANOVA.

The effect of TIME simply represents the difference between acquiring the information directly and recalling it several weeks later. It is, therefore, no surprise that the percentage of correct answers was higher during Test 1 (mean ± SD = 84 ± 14%) than during Test 2 (mean ± SD = 49 ± 20%).

The main effect of METHOD shows that the participants learning with the VR atlas achieved better results than those learning with books. The post hoc t-tests showed that, more specifically, the results of the OB and VR groups were significantly different in the second test, not in the first. The improvement in test results can thus be attributed to better memory retention in the VR group; this confirms hypothesis 2.

Taken together with the response time effect described above, this study has shown that, under the given conditions, the VR atlas both enabled faster information acquisition and facilitated improved memory retention. This makes the immersive anatomy atlas an overall more efficient tool for learning anatomical knowledge than classical learning through books. Combining the active learning and exploration already possible in the VR atlas with additional methods like tests and gamification, which can be added to VR comparatively easily, might enhance the performance of VR learning even further. Additionally, the constructivist learning aspects already present in VR environments could be strengthened with further technological additions. Presently, the virtual atlas allows for self-guided exploration in a relevant and realistic environment, and enables the learners to take ownership of their learning (see for aspects of constructivist learning; Amarin and Ghishan, 2013; Johnson-Glenberg, 2018). Future developments of the VR-atlas may allow multiple people to enter the same simulation, adding social interaction and collaboration (Amarin and Ghishan, 2013), and may add the option for users to construct the virtual environments themselves, which improves learning success especially for low-performance students (Winn et al., 1997).

In summary, the immersive anatomy atlas already seems more efficient than classical learning modalities, and future

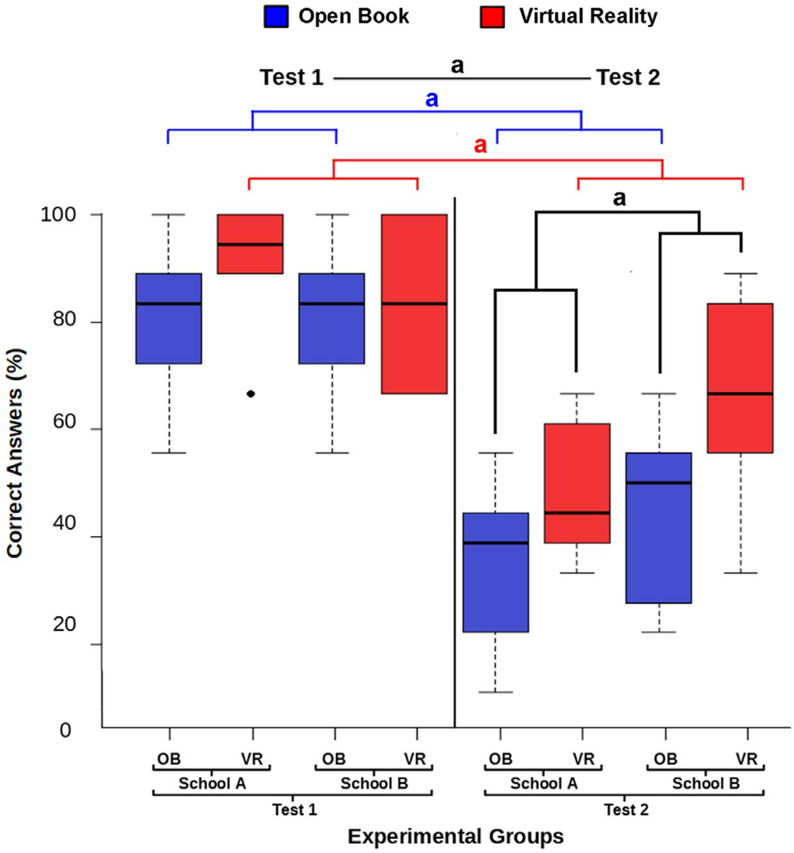

Figure 5.

Box plots depicting the percentage of correct answers for schools A and B, in Test 1 (first test, learning) and Test 2 (second test, weeks later, recognition/ recall), for both OB and VR. The black dot denotes an outlier. There was a significant main effect of TIME (Test 1 vs. Test 2), a significant main effect of learning METHOD (OB vs. VR), and a significant interaction effect of TIME and SCHOOL (School A, Test 2 vs. School B, Test 2). OB, open book; VR, virtual reality. [a]Significant differences between groups ($P < 0.01$). -

developments of VR in general and this software in particular, are expected to increase this advantage.

The significant interaction effect between SCHOOL and TIME is an ordinal effect for TIME; The mean percentage of correct answers is consistently lower in Test 2 than in Test 1. This means that the global main effect of TIME reported above is unaffected by the interaction effect.

The associated post hoc t-tests produced five significant differences. Four of those, however, contained the TIME effect and thus offer no new insight. The last was the significant difference between (school A, Test 2) and (school B, Test 2). Apparently, the retention of anatomical knowledge was overall better in school B (mean ± SD = 55 ± 21%) than in school A (mean ± SD = 42 ± 16%). However, the reason for that effect could also be the different extent of time between the tests in the two schools.

The anatomy tests employed in this study (see Supplemental Material File) to assess the learned and retained knowledge seemed to have performed well enough, but could be improved in future studies. The questions used in Test 1 were of comparable difficulty; The difficulty index for this test was 83.68 (±13.82), with a range of 68.75%–100%. Question 1 seemed to suffer from a ceiling effect as participants from both conditions were able to answer it with average correctness of 100%. This test fulfilled its core role to teach the relevant knowledge and give room to engage with the teaching material, but was sub-optimal in differentiating high- and low-performing students. Test 2 was overall more difficult, which is of course due to the time elapsed between learning and recall. Aside from that, the variability in difficulty between questions was also larger than in Test 1, with a difficulty index of 48.61 (±19.71) and a range of 9.38–100. Questions 1 and 6 seemed to have been too difficult. Future studies should substitute them with easier alternatives and use question 5, the easiest, at the start of the test.

The participants in this study had no previous knowledge in the field of medicine, which makes the generalization to more experienced medical personnel difficult. Previous literature for a population of experienced medical students had suggested a lack of advantages of immersive VR (Stepan et al., 2017). However, this study has shown the usefulness of VR environments for the initial acquisition of anatomical knowledge, which is in accordance with the findings of Zinchenko and colleagues (Zinchenko et al., 2020).

Virtual reality seems on its way to becoming an integral part of education and training for medical vocations (Rizzetto et al., 2020), and this study supports this direction. Future research should focus on the use of tools like the immersive anatomy atlas for medical trainees or students, especially in the early stages of teaching. If the VR application has a high physical fidelity, inexperienced students may gain more from its use (Birbara and Pather, 2021), and this learning experience could prepare students for the eventual cadaver dissection courses. A similar statement about the usefulness of VR early in teaching has been made by Andersen and colleagues regarding the acquisition of surgical skills, after they found that for novices cognitive load is higher during cadaver training than VR training (Andersen et al., 2016). In general, supplementing cadaver courses with VR applications for preparation and repetition makes sense, given some of the problems cadaver studies face, like high financial expenses, limited availability, or high student to cadaver ratios (Wainman et al., 2021).

Limitations of the Study

Some limitations should be considered in regard to this study and its results.

The two schools had different time intervals between the initial learning and the test for knowledge retention. Any difference between the schools could thus be attributed to that discrepancy. Therefore, any further in-depth comparison of the schools and how their curricula may have affected the results became impossible.

The questions used to test the participants' anatomical knowledge were not from a standardized questionnaire. Instead, they were specifically created for this study. This had the advantage of being a perfect fit for the purpose of the experiment, but the disadvantage was that the tests lacked established values for quality criteria such as reliability and validity.

This study compared immersive VR to anatomy textbooks; There are, however, other ways to learn human anatomy. Especially physical models, which have been shown to be superior to VR in some contexts (Wainman et al., 2020), could have been included as a third group to provide a more complete overview in this article.

Lastly, it has to be noted that the sample size was small; with only eight participants per group per school, the results have to be viewed with caution.

CONCLUSIONS

Based on the results of this study, immersive VR learning seems to be more efficient and to facilitate better long-term retention of knowledge in previously inexperienced students. The reason for that could lie in the VR environment's high immersion and the possibility to freely explore a realistic replication of human anatomy. There are many possibilities for medical teaching and training which VR technology offers, the extent of which might grow with the advancements of the hard- and software. A future challenge for anatomical educational research will be establishing a meaningful standard for a curriculum which combines immersive VR, classic textbooks, and cadaver training.

ACKNOWLEDGMENTS

The authors thank the students of high schools A and B for participating in the experiment and thereby supporting this research. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

NOTES ON CONTRIBUTORS

KILIAN GLOY, Dr. Rer. Nat., is a research fellow at the Department for Visceral Surgery at the University Clinic Oldenburg, Carl von Ossietzky University Oldenburg in Oldenburg, Germany. His research interest is in human learning and decision making, as well as the neuropsychological foundations thereof.

PAUL WEYHE, is a twelfth-grade high school student who worked as a high-school research assistant in cooperation with the Department for Visceral Surgery at the University Clinic Oldenburg, Carl von Ossietzky University Oldenburg in Oldenburg, Germany. He has a keen interest in science.

ERIC NERENZ, is a twelfth-grade high school student who worked as a high-school research assistant in cooperation with the Department for Visceral Surgery at the University Clinic Oldenburg, Carl von Ossietzky University Oldenburg in Oldenburg, Germany. He has a keen interest in science.

MAXIMILIAN KALUSCHKE, M.Sc., is a researcher at the faculty for Mathematics and Computer Science at the University of Bremen in Bremen, Germany. His research interests are collision detection, haptic rendering, and physically based simulation.

VERENA USLAR, Dr. Rer. Nat., is a senior research scientist in the Department for Visceral Surgery, at University Hospital for Visceral Surgery, the Carl von Ossietzky University Oldenburg in Oldenburg, Germany. She teaches all aspects of good scientific practice to medical students and clinicians and her research interest is in the use of new technologies in the perioperative setting.

GABRIEL ZACHMANN, Ph.D., is a professor for computer graphics and virtual reality in the Institute for Computer Graphics and Virtual Reality at the University of Bremen in Bremen, Germany. He teaches virtual reality, computer graphics, and geometric computing to Bachelor's and Master's students. His research interests include geometric algorithms for computer graphics, massively parallel algorithms on the GPU, virtual prototyping, and medical immersive simulators.

DIRK WEYHE, M.D., is a professor of visceral surgery and Head of the Department for Visceral Surgery at the University Clinic Oldenburg, Carl von Ossietzky University Oldenburg in Oldenburg, Germany. His research interests are in digital medicine and carcinoma studies.

LITERATURE CITED

Amarin NZ, Ghishan RI. 2013. Learning with technology from a constructivist point of view. Int J Bus Humanit Technol 3:52–57.

Andersen SA, Mikkelsen PT, Konge L, Cayé-Thomasen P, Sørensen MS. 2016. Cognitive load in mastoidectomy skills training: Virtual reality simulation and traditional dissection compared. J Surg Educ 73:45–50.

Birbara NS, Pather N. 2021. Real or not real: The impact of the physical fidelity of virtual learning resources on learning anatomy. Anat Sci Educ 14:774–787.

Birbara NS, Sammut C, Pather N. 2020. Virtual reality in anatomy: A pilot study evaluating different delivery modalities. Anat Sci Educ 13:445–457.

Bracq MS, Michinov E, Arnaldi B, Caillaud B, Gibaud B, Gouranton V, Jannin P. 2019. Learning procedural skills with a virtual reality simulator: An acceptability study. Nurse Educ Today 79:153–160.

Brewer-Deluce D, Bak AB, Simms AJ, Sinha S, Mitchell JP, Shin D, Saraco AN, Wainman BC. 2021. Virtual reality bell-ringer: The development and testing of a stereoscopic application for human gross anatomy. Anat Sci Educ 14:330–341.

Dua A, Coppola KM, Mulheron GW, Troupe D, Lebeau R. 2021. Development of a novel peer-sharing application to supplement learning from cadaveric dissection. Anat Sci Educ 14:491–504.

Field A, Miles J, Field Z. 2012. Discovering Statistics Using R. 1st Ed. London, UK: SAGE Publications Ltd. 992 p.

Fosnot CT, Perry RS. 1996. Constructivism: A psychological theory of learning. In: Fosnot CT (Editor). Constructivism: Theory, Perspectives, and Practice. 1st Ed. New York, NY: Teachers College Press. p 8–33.

Frederiksen JG, Sørensen SM, Konge L, Svendsen MB, Nobel-Jørgensen M, Bjerrum F, Andersen SA. 2020. Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: A randomized trial. Surg Endosc 34:1244–1252.

Freina L, Ott M. 2015. A literature review on immersive virtual reality in education: State of the art and perspectives. In: Proceedings of the 11th International Conference eLearning and Software for Education (eLSE 2015): Rethinking Education by Leveraging the eLearning Pillar of the Digital Agenda for Europe; Bucharest, Romania, 2015 April 23-24. p 133–141. Carol I National Defense University, Bucharest, Romania.

Gureckis TM, Markant DB. 2012. Self-directed learning: A cognitive and computational perspective. Perspect Psychol Sci 7:464–481.

Hazlett CB. 2009. Prerequisite for enhancing student learning outcomes in medical education. Sultan Qaboos Univ Med J 9:119–123.

Huber T, Paschold M, Hansen C, Wunderling T, Lang H, Kneist W. 2017. New dimensions in surgical training: Immersive virtual reality laparoscopic simulation exhilarates surgical staff. Surg Endosc 31:4472–4477.

Johnson-Glenberg MC. 2018. Immersive VR and education: Embodied design principles that include gesture and hand controls. Front Robot AI 5:81.

Koivisto J, Hamari J. 2019. The rise of motivational information systems: A review of gamification research. Int J Inform Manag 45:191–210.

Kornell N, Hays MJ, Bjork RA. 2009. Unsuccessful retrieval attempts enhance subsequent learning. J Exp Psychol Learn Mem Cognit 35:989–998.

Krathwohl DR. 2002. A revision of Bloom's taxonomy: An overview. Theor Pract 41:212–218.

Krokos E, Plaisant C, Varshney A. 2019. Virtual memory palaces: Immersion aids recall. Virtual Reality 23:1–15.

Kurul R, Ögün MN, Neriman Narin A, Avci S, Yazgan B. 2020. An alternative method for anatomy training: Immersive virtual reality. Anat Sci Educ 13: 648–656.

Markant DB, Ruggeri A, Gureckis TM, Xu F. 2016. Enhanced memory as a common effect of active learning. Mind Brain Educ 10:142–152.

Parong J, Mayer RE. 2018. Learning science in immersive virtual reality. J Educ Psychol 110:785–797.

Paulsen F, Waschke J. 2017a. Sobotta, Atlas der Anatomie Band 1: Allgemeine Anatomie und Bewegungsapparat. 24th Ed. Munich, Germany: Elsevier GmbH / Urban & Fischer. 469 p.

Paulsen F, Waschke J. 2017b. Sobotta, Atlas der Anatomie Band 2: Innere Organe. 24th Ed. Munich, Germany: Elsevier GmbH / Urban & Fischer. 352 p.

Paulsen F, Waschke J. 2017c. Sobotta, Atlas der Anatomie des Menschen Band 3: Kopf, Hals und Neuroanatomie. 24th Ed. Munich, Germany: Elsevier GmbH / Urban & Fischer. 941 p.

Pius-Hospital Oldenburg. 2018. Immersive virtual reality anatomy atlas V02. YouTube video. Pius-Hospital Oldenburg, Oldenburg, Germany. 1:24 min. URL: https://www.youtube.com/watch?v=JY50Wjh-olw [accessed 23 November 2020].

Rizzetto F, Bernareggi A, Rantas S, Vanzulli A, Vertemati M. 2020. Immersive virtual reality in surgery and medical education: Diving into the future. Am J Surg 220:856–857.

Sankaranarayanan G, Odlozil CA, Wells KO, Leeds SG, Chauhan S, Fleshman JW, Jones DB, De S. 2020. Training with cognitive load improves performance under similar conditions in a real surgical task. Am J Surg 220:620–629.

Selzer MN, Gazcon NF, Larrea ML. 2019. Effects of virtual presence and learning outcome using low-end virtual reality systems. Displays 59:9–15.

Seymour NE. 2008. VR to OR: A review of the evidence that virtual reality simulation improves operating room performance. World J Surg 32:182–188.

Siemens G. 2005. Connectivism: A learning theory for the digital age. Int J Instr Technol Dist Learn 2:1.

Stepan K, Zeiger J, Hanchuk S, Del Signore A, Shrivastava R, Govindaraj S, Iloreta A. 2017. Immersive virtual reality as a teaching tool for neuroanatomy. Int Forum Allergy Rhinol 7:1006–1013.

Wainman B, Aggarwal A, Birk SK, Gill JS, Hass KS, Fenesi B. 2021. Virtual dissection: An interactive anatomy learning tool. Anat Sci Educ 14:788–798.

Wainman B, Pukas G, Wolak L, Mohanraj S, Lamb J, Norman GR. 2020. The critical role of stereopsis in virtual and mixed reality learning environments. Anat Sci Educ 13:401–412.

Weyhe D, Uslar V, Weyhe F, Kaluschke M, Zachmann G. 2018. Immersive anatomy atlas—Empirical study investigating the usability of a virtual reality environment as a learning tool for anatomy. Front Surg 5:73.

Winn W, Hoffman H, Hollander A, Osberg K, Rose H, Char P. 1997. The effect of student construction of virtual environments on the performance of high-and low-ability students. In: Proceedings of the Annual Meeting of the American Educational Research Association (AERA 1997): Talking Together in Educational Research and Practice; Chicago, IL, 1997 March 24-28. p 1–7. American Educational Research Association, Washington, DC.

Wong A, Leahy W, Marcus N, Sweller J. 2012. Cognitive load theory, the transient information effect and e-learning. Learn Instr 22:449–457.

Yeh DD, Park YS. 2015. Improving learning efficiency of factual knowledge in medical education. J Surg Educ 72:882–889.

Yildirim C. 2019. Cybersickness during VR gaming undermines game enjoyment: A mediation model. Displays 59:35–43.

Zackoff MW, Real FJ, Cruse B, Davis D, Klein M. 2019. Medical student perspectives on the use of immersive virtual reality for clinical assessment training. Acad Pediatr 19:849–851.

Zinchenko YP, Khoroshikh PP, Sergievich AA, Smirnov AS, Tumyalis AV, Kovalev AI, Gutnikov SA, Golokhvast KS. 2020. Virtual reality is more efficient in learning human heart anatomy especially for subjects with low baseline knowledge. New Ideas Psychol 59:100786.