Anatomy Education Applied to Dental Surgery in Virtual Reality

Bernardo Gonçalves de Faria
Instituto Superior Técnico, Portugal
bernardo.faria@tecnico.ulisboa.pt

ABSTRACT

Dental anatomy is currently taught and studied using two methodologies: anatomical atlases and cadaveric material. The former is known to be quite laden with terms and concepts; the latter needs donations and unique storage conditions for its preservation. With the development of VR technology, these two paths can merge into one. Our work presents IMPLANTIGRAPH, a study and teaching tool that aims to explore the use of anatomical graphs and 3D labeling, containing anatomical information focused on Implantology, in an immersive environment, using virtual reality glasses and the respective controllers that allow the user to interact with any of the chosen information designs. We evaluated IMPLANTIGRAPH through anatomical questionnaires and subjective metrics with 30 Master's students in Dentistry and 3 anatomy teachers. The results showed that the use of anatomical graphs in an immersive environment, despite not being quite appealing, facilitates the process of learning and teaching anatomy applied to Implantology, thus benefiting the students' study and the teachers' teaching methodology, but only as a complementary tool to conventional methods. Hence, IMPLANTIGRAPH proved to be a promising tool for anatomy education applied to Implantology and a starting point for exploring the concept of anatomical graphs in virtual environments.

CCS CONCEPTS

• **Applied computing** → **Interactive learning environments**; • **Human-centered computing** → **Virtual reality**; **User studies**; Interface design prototyping; User interface design;

KEYWORDS

Anatomy Education, Virtual Reality, Dental Implantology, Graphs, Knowledge Representation, Surgical Anatomy

1 INTRODUCTION

Anatomy manuals and cadaverous materials are the key methods for studying and teaching anatomy. Manuals come with medical illustrations populated with labels [25, 31, 34, 36]. This visual mapping is highly relevant but it is challenging for students to process and memorize such a massive amount of information [29, 33]. Moreover, conventional labeling requires students to perform the cognitive load to build a mind map [12, 17, 33]. Cadaverous material has the advantage of spatial visualization but hides its continuous maintenance to avoid degradation, and donations are quite rare. All of this makes the study of anatomy a great challenge for students [18, 25, 26, 31, 39, 41].

The textual information can be presented as knowledge representations in the form of acyclic anatomy graphs, creating an alternative method for memorization [12]. From "general to particular" reasoning, the main concept becomes the root of the map, which branches into its several components, and these branch into others, and so forth, generating an easy-to-read hierarchy that helps the topographic perception of anatomical structures [7]. This technique was widely studied by Tony Buzan (which he named *Mind Maps* [11]). According to Buzan, the use of mind maps as a visual map can increase memory retention and facilitate the understanding of concepts and information.

Several studies explored the idea of using graphs as mental models [7, 10, 12, 16]. However, these examples do not use acyclic graphs, do not fully represent Buzan's technique, and are not centered on dental implantology. Nonetheless, they demonstrate the power mental models have to facilitate understanding and retention.

With the increasing development of VR technology, cadaverous material can be simulated in immersive VR educational tools, leading to more interesting and fashionable learning methodologies [26, 27, 41]. Studies show that using VR in anatomy improves the anatomical learning effect on medical students and that there has been an increase in adapting to digital modalities in anatomy education [5, 19, 21, 26]. Regarding dental anatomy specifically, studies show that dental students using VR technologies can learn motor skills more quickly compared to students using only phantom training [13], the use of VR technology is useful both in teaching as well as in the learning process [22, 27, 40], by giving students a better understanding of the subject being taught and a more effective knowledge transfer and retaining, both in short-term and long-term application [21, 24].

A variety of VR applications regarding anatomy have been developed: some works explore general anatomy study [18, 37, 41]; other works are focused on dental medical procedures, however, they do not have a theoretical component [20, 35, 42, 43]. There is even one study that focuses on the reconstruction of dental cavities but uses Optical See-Through AR and not VR [19].

In this work, we want to evaluate the potential of using 3D immersive anatomical graphs in the learning and teaching process of topographic anatomy applied to dental implantology, adapting conventional labeling in virtual environments complemented with anatomy graphs. Moreover, topographic anatomy is based on anatomical regions [4], therefore we selected the more important head anatomy concepts by each region.

As such, the research questions of this work are (1) Can knowledge representation as anatomy graphs inside VR educational tools for oral surgery facilitate the learning and perception of anatomical structures and understanding of its topology relations?, (2) Can knowledge representation as anatomy graphs inside VR educational tools benefit oral surgery education?, and (3) Are anatomy graphs inside VR educational tools a better approach to teach oral surgery, rather than current approaches such as visual textbooks?

To help answer the previous questions, we developed a high-fidelity VR prototype with different layouts representing dental anatomical concepts. We want to verify the hypotheses (1) Anatomy graphs in VR environments help dental students to perceive mental models faster than current learning methods, and (2) Anatomy graphs in VR environments benefit the learning and teaching of dental surgical anatomy.

In essence, the goals and contributions of our work can be summarized as (1) IMPLANTIGRAPH, a VR educational tool aimed to assist anatomy learning and teaching in dental surgery, focusing on the field of implantology. The design process brought together interviews with 2 dentistry teachers and 16 Master's students who have attended anatomy classes in the past and co-design sessions with a dentist and dentistry teacher; (2) We report a comparative study between the use of the conventional training method applied to VR, and the use of anatomy graphs together with the conventional layout, both implemented in IMPLANTIGRAPH. 30 students and 3 teachers, were recruited to participate in the user study, where we measure the viability of the prototype through anatomical quizzes, system usability, workload measures, sense of presence, user satisfaction, and preferences; and (3) We gathered more feedback through semi-structured interviews with all the participants regarding their experience with IMPLANTIGRAPH.

2 RELATED WORK

In this section, we present examples of works that explore the use of graphs or mind maps as knowledge representation, as well as some examples of past VR applications dedicated to general anatomy and dental anatomy.

2.1 Mind Maps / Graphs

Daeijavad et al. [14] questioned how to display and organize 3D data information in multiple views inside an immersive environment, so they developed a taxonomy for multiple immersive layouts by creating several of them on a 3D space. Although they did not test complex information, they conclude that the more data, the more difficult is to design the layouts. From here, we followed a guideline on how to present and place a panel of static information in the user's Content Zone, ending up choosing a flat panel that can be tilted or rotated by the user.

Wen et al. [38] ask the same question because multiple display views of large amounts of data are normally presented on 2D visualization displays. They explored the design space of multiple-view representations in immersive environments by examining their effects on situated analytics to achieve the best of both worlds (high situatedness and effective analytics) in an automatic layout adaptation prototype developed by the team. The resulting layouts facilitate the users in completing several tasks.Due to the limited FOV in their HDM they could only test up to six views; moreover, all views tested were static and had a fixed size. From here, we also followed a guideline on how to present and place the floating tags surrounding the 3D models.

Tony Buzan widely studied the idea of graphs as a mind map. Figure 1 presents an example of a hand-drawn mind map of the Posterior Mandible, based on the ones designed by Tony Buzan.

2.2 Mind Maps / Graphs in Anatomy

The idea of using graph-like diagrams in anatomy has been explored for a couple of years. One of the first digital knowledge representations concerning anatomy was developed by Schubert et al. [32]

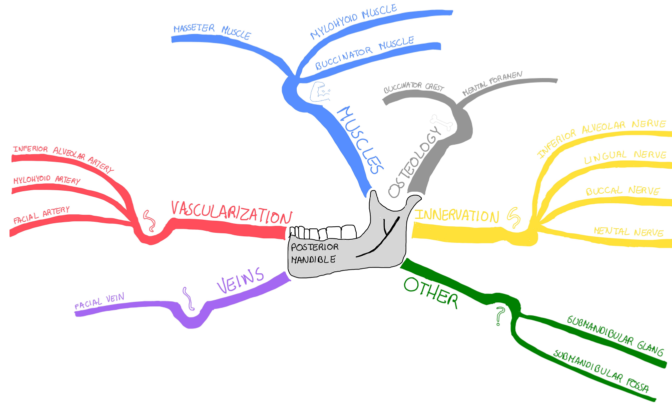

Figure 1: Example of a hand-drawn mind map, based on Tony Buzan's, within the context of this work.

in the form of a semantic net. Combining computer graphics with knowledge engineering, a new approach to knowledge representation using a volume-based data structure was developed, using three-dimensional visualization of anatomical concepts, allowing for several domains of representations.

More recent work demonstrates that this technique continues to be explored today. Anatomical variations and their occurrence frequencies are essential to correctly diagnose and safely treat patients. Currently, this information is presented in textual information, requiring the reader to construct a mental model, which becomes more challenging with more complexity. Smit et al. [33] developed VarVis, an interactive visualization application for anatomical variations to compare and explore variations on branching structures at a local or global level, using illustrations of variations to create graphs that define nodes for every endpoint and junction.

2.3 VR in Anatomy

VR in medical education not only increases physical interaction but aids in information recall. For anatomy courses, a VR educational tool could be effective in learning and retaining anatomy knowledge besides textbooks. Gloy et al. [18] developed an immersive and interactive 3D anatomy atlas to freely explore anatomy structures of the human body through virtual dissection, fastening the acquisition of new information and improving the retention of knowledge. The atlas uses a head-mounted display and controllers to interact with the environment. The atlas was tested with non-medical students and results proved that by using VR, acquiring unknown information is faster and memory retention is improved.

Another work developed in VR and focusing on anatomy was tested by Yamazaki et at. [39]. In this case, VR was explored as a substitute to train surgical procedures on bones, since 2D computer screens and cadaver bone drilling offer limited resources. Yamazaki created a 3D model of temporal bones to be used on a head-mounted display to demonstrate the application of VR in preoperative planning and usage. CBCT images were used to create the 3D models, and a session to first manipulate the bone and second to fill a questionnaire to assess its validity was conducted. Most participants were favorable about using the VR model, considering the VR technology superior to a 2D screen for both the training procedure and using it as an educational tool.

2.4 VR in Dental Anatomy

VR is being specifically explored in dental anatomy. As known, dental anesthesia is a challenging clinical procedure to master, so every dentist must be competent in doing it, and training on a plastic manikin head does not provide reliable feedback. Grandhi et al. [20] developed a VR-based system to train dental anesthesia by giving users visual, auditory, and haptic feedback. The tool consists of a 3D head model from a real patient, virtual hands to interact with the environment, a control pad to perform real physical adjustments, sound and haptic feedback, two modes of usage (Practice and Assessment) and error counting. A user study was conducted, revealing the tool to be useful when integrated into the pre-clinical curriculum. For future work, the authors stated multi-point haptic feedback and multi-user ability.

Another challenge for dental students is to master the procedure of performing a dental implant. The shortage of this training affects their performance, resulting in a lack of precision and inadequate implant placement. Zorzal et al. [43] developed IMMPLANT, a virtual reality educational tool to assist implant placement learning, helping students by manipulating 3D dental models. The system was tested and the results showed that the application constitutes a versatile and complementary tool to assist implant placement learning by promoting immersive visualization and spatial manipulation of 3D dental anatomy.

2.5 Limitations and Gaps

Regarding mind maps and graphs, Daeijavad et al. [14] do not test the design choices for layout performance on concrete examples; the system of Wen et al. [38] do not let users select views of interest or move the information, and the views are always static with a fixed size; Brinkley et al. [9] referred how time-consuming it is to segment CT medical images and lacks anatomical concepts; at last, the work of Smit et al. [33] lacks a larger study with medical students, since they only tested three participants.

Regarding VR, Gloy et al. [18] referred that the questions used in the study were specifically created for the study ("a perfect fit"), and lacked a comparison between the VR application with other ways of learning anatomy; the system of Yamazaki et al. [39] does not offer real-time feedback on their application; the work of Grandhi et al. [20] does not have multi-point haptic feedback integration, no multi-user ability, and they did not perform a pilot test of the system; Zorzal et al. [43] warns that wearing a VR headset can force the users to continuously change their posture, therefore affecting their performance, and do not evaluate cognitive load.

From the previous limitations and gaps, our work addresses some of them, while others were already thought of as future work. Our system is divided into four activities, with each activity having its own content that is shown on two different layouts; therefore, we test the same content on different layouts per activity, the size of the layouts is not fixed (some activities have more concepts than others), and both layouts are movable in space (not static). Both layouts are tested on a 3D immersive environment with a generous number of medical students and teachers as participants. At last, our system was designed to be used while sitting down, therefore there is no need for rapid or prolonged movements that force participants to change their posture.

3 REQUIREMENTS ELICITATION

3.1 Interview Sessions

We handled sessions over two days, with 2 anatomy teachers with ages 58 and 61, with Ph.D. Degrees in Dentistry and 16 students currently taking a Master's Degree in Dentistry, aged between 21 and 29. Both teachers were male; 6 students were male and 10 were female. They were asked to fill in an Informed Consent form and Demographic Profile form in order to participate in semi-structured interviews. These interviews were helpful in gathering user requirements and needs about the teacher's methodology in lecturing and evaluating the students, and the dental student's methodology in their own study and their feedback regarding anatomy classes. The interviews were audio recorded with all the participant's consent.

3.1.1 ***Teachers' Interviews.*** The teachers' interview was divided into four different groups of questions: *Conventional Teaching and Learning Tasks*, *"Out Of The Box" Tasks*, *Knowledge Assessment Methods*, and *Virtual Reality as a Learning and Education Method*. For the first two groups, the teachers responded with the most common methods, such as PowerPoint presentations, diagrams, videos, and physical models, but both answered that do not use "out of the box" methods to teach. One teacher manifested curiosity in using 3D technology to improve the learning tradition. Regarding *Knowledge Assessment Methods*, the evaluation of the students is conducted by oral and written assessments. For the final set of questions, both teachers express their motivation to use VR tools in the teaching process, yet none of them ever used such technology.

3.1.2 ***Students' Interviews.*** The students' interview was divided into two groups: *Study and Learning Methods*, and *Lecturing Methods*. The first group gathered information regarding the student's study methods, like PowerPoints, textbooks, pictures, online material, and even hand-written or hand-drawn sketches provided later by some of the students (Figure 2), as well as limitations and difficulties while studying anatomy (a vast number of concepts, making it difficult to memorize). The last group gathered feedback from problems students have with attending anatomy lectures, such as PowerPoints with a lot of textual content instead of visual information; for practical classes, the main problem was too much detail to memorize. Feedback for changes was also inquired: applications with 3D models, more videos, and physical models for theoretical classes; more videos and applications with 3D models for practical classes.

One student mentioned the use of *mind maps* as a form of studying and memorizing. She started using this technique in secondary school to help her internalize concepts faster and better understand hierarchies, and took this method to university, using it extensively in anatomy classes to memorize theoretical content. She drew the maps by hand, like a spider: the body was the main concept, and the various legs formed the "first level of a hierarchy", which could in turn branch out into other legs, and so on. Unfortunately, she does not have photographs or examples of maps made by herself because she passed this study material on to other colleagues, who

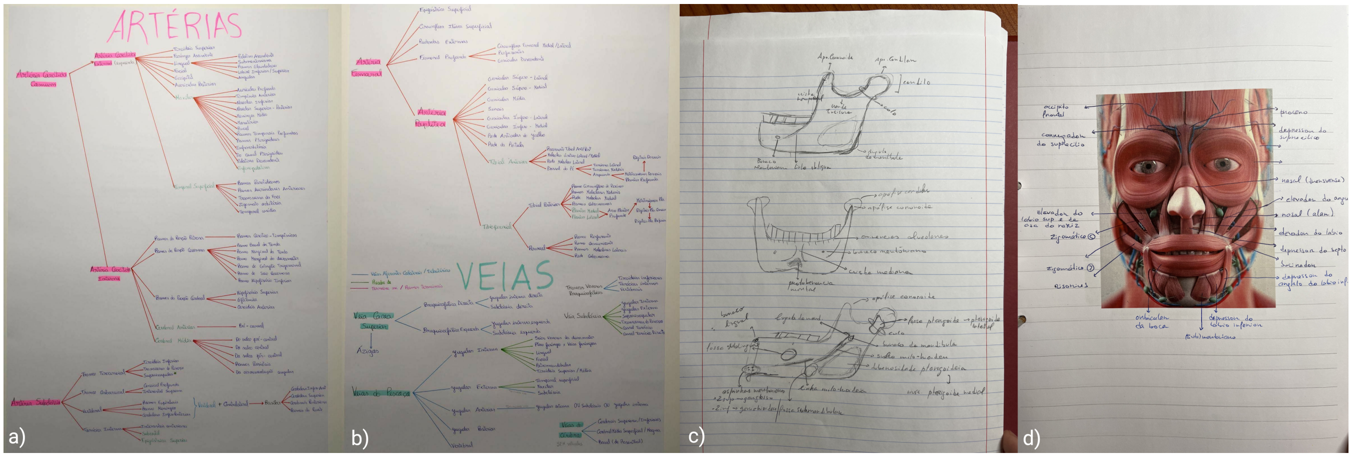

Figure 2: Some students' hand sketches: a) and b) arteries and veins, c) bones of the mandible, and d) muscles of the face.

passed it on to others, from which we concluded that this method of study works well for anatomical study. The technique used by this student adapts to the theory of *mind maps* created by Tony Buzan, mentioned in the Introduction, thus strengthening the relevance of mind maps applied to the study of anatomy.

3.2 Co-Design Sessions

Based on the interview sessions, several co-design sessions with a dentistry teacher were relevant to obtain some medical data to be included in the initial prototype. We divided the maxilla and the mandible into six different regions, with anatomical concepts relevant to each region. From an open-access dataset of patient-specific human jaw models, the three most completed were chosen. The anatomic information was then hand-sketched by region on each of the jaw models to formulate the 3D conventional layout, resulting in three low-fidelity prototypes. A consensus was later made to elect the most complete and correct model to use on the final prototype. Figure 3 represents the first anatomical concepts of one region hand sketched on the chosen model. Then it was time to design the other layout, an acyclic graph, to the side, of the entire constitution of each region. Figure 4 shows the design of an entire view of one region.

A mid-fidelity prototype started to be developed on the Unity engine, consisting of the Premolar Mandibular region represented on a 3D model of the mandible with floating tags associated with colored buttons (the conventional layout). Each button enables or disables a concept, and each color represents the different types of constituents: grey represents Osteology, blue represents Muscles, red represents Vascularization, purple represents Veins, yellow represents Innervation, and green represents relevant elements that do not fit into the classifications described before. Feedback from the co-design participant turned out positive. The professional agreed that it was useful to have colors on the buttons because it helps to identify the type of elements being pointed out. One suggestion to be considered was the rotation and spatial movement of the 3D model so that the user could rotate it, being able to see the entire model and also decide on the distance to the camera. The anatomic regions were discussed again and, for the sake of simplicity, were reduced to just four: anterior maxilla, posterior maxilla, anterior mandible, and posterior mandible.

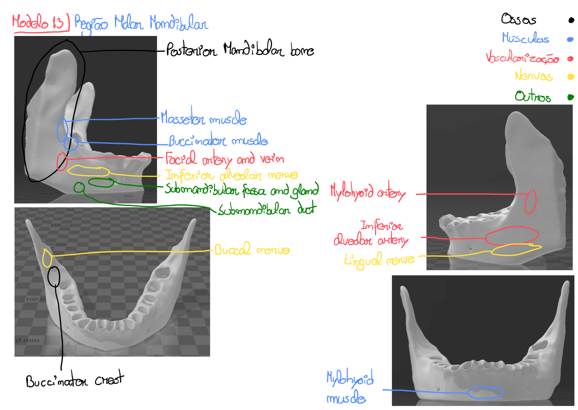

Figure 3: First concepts hand sketched on the Posterior Mandible region on the most complete model.

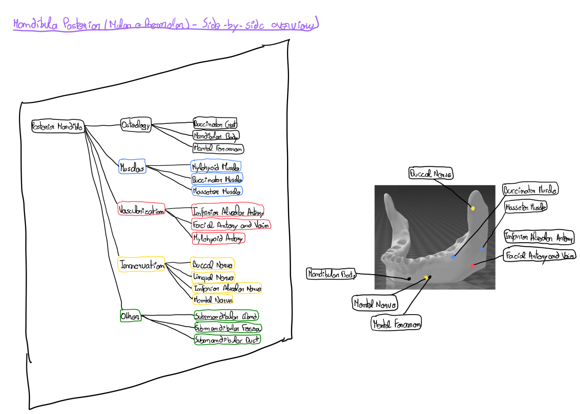

Figure 4: Design of the two layouts of one region: the conventional layout on the right and the graph to the side on the left.

4 IMPLANTIGRAPH

We developed a high-fidelity prototype, called IMPLANTIGRAPH, to address the benefits and limitations of VR anatomy maps.

4.1 3D Models

To obtain a patient-specific model of an entire jaw, it was necessary to extract the 3D model from full-toothless jaw CBCT images provided by Cooperativa de Ensino Superior Egas Moniz. The 3D reconstruction process followed the pipeline described by Paulo et al. [30]. However, the model resulting from the above procedure was considered invalid for anatomical study, as the patient in question did not have enough bone on the anterior maxilla and mandible. A repository containing an open-access dataset of 17 patient-specific textureless STL models of human jaws was found, each one divided into a mandible model and a maxilla model. The most complete model was chosen. To add realism, a procedural bone texture was created using the Blender software (version 3.3. LTS), and the model was exported as an FBX file and imported into the Unity project.

4.2 Apparatus

Our application was developed in Unity3D version 2021.3.8f1. During development, the HMD in use was the Oculus Quest 2, and for the user study, the Oculus Quest 1 was the chosen one.

4.3 Interaction

To interact with the 3D models, the user must use the left thumbstick for spatial movement (translation) and the right thumbstick for horizontal or vertical rotation.

The floating tags, from now on **floating graph**, consist of clickable buttons with different colors and layouts. The triggers on the Oculus controllers activate or deactivate the buttons. The floating graph is linked to the 3D model, meaning that whenever the model is moved in space, the floating graph follows its movement. Figure 5 presents the floating graph of the Posterior Mandible.

The graph to the side, from now on **side graph**, is a dark-background plane with multiple colored boxes organized to create a visual hierarchy. Starting at the main structure, it branches into the main divisions (Osteology, Muscles, etc.), and the latter are subdivided into the different anatomical constituents to be studied. The graph is grabbable through the grab button, the user can move it around and place it where it is most convenient. When clicking on the leaf nodes using the trigger button, the node highlights and the associated floating tag is activated, herein the link between both layouts. Figure 6 presents the side graph of the posterior mandible.

5 USER STUDY

A user study was performed to evaluate the benefits of using VR anatomy graphs as a studying and teaching tool.

5.1 Participants

A total of 33 participants were invited to take part in our user study: 30 Master's students in Dentistry, and 3 teachers. The students (20 female, 10 male), with ages ranging from 21 to 31 (Mean = 23.5, Standard Deviation (SD) = 2.7), were all perceiving a Master Degree in Dentistry, with one of them being employed as an Oral Hygienist for 5 years. Of the 30 students, 9 of them referred that they never dealt with virtual reality technology. As for the 3 teachers (all male), with ages ranging from 25 to 51 (Mean = 34.7, SD = 11.6), two have a Post-Graduate Degree in Dentistry and one a Ph.D. Degree in Dentistry. They are all employed with specialties in *Implantology*, *Oral Rehabilitation and Implantology*, and *Oral Pathology and Surgery (with a subspecialty in Oral Cancer)*, and the years of experience range between 2 and 25 (Mean = 10.7, SD = 10.2). One of the professors never dealt with virtual reality technology.

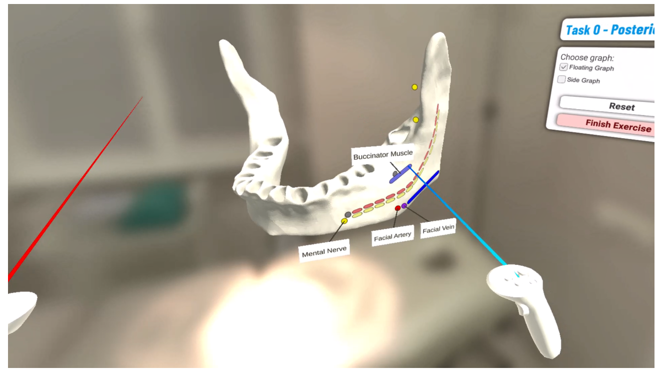

Figure 5: The floating graph of the Posterior Mandible.

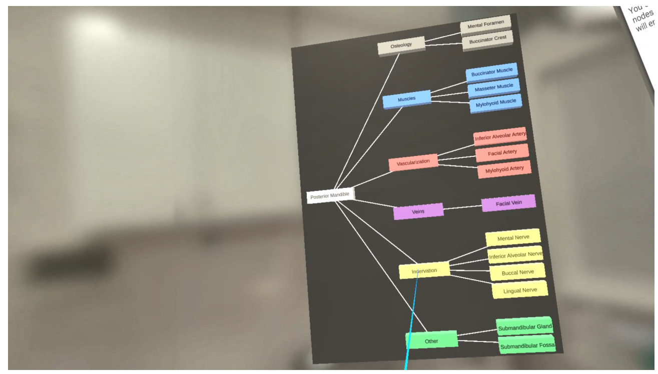

Figure 6: The side graph of the Posterior Mandible.

5.2 Apparatus

The user study took place at Cooperativa de Ensino Superior Egas Moniz. The setup consisted of an Oculus Quest 1 headset, the IMPLANTIGRAPH prototype loaded on the headset, and two portable computers: the first to support the prototype casting of the Oculus and the second to fill in the questionnaires and answer the quizzes.

5.3 Variables

One independent variable was chosen to evaluate the labeling method, with values "conventional labeling" (the floating graph) and "side-by-side labeling" (the side graph). The dependent variables were divided between objective measures (time of each quiz, accuracy of each quiz, and number of concepts) and subjective measures (participant preferences for the use of a VR tool, together with the labeling strategies, and the use of anatomic graphs in a virtual environment).

5.4 Tasks

The evaluation method for the Master's students used the between-group design. Students were divided into two groups: Group A used only the floating graph, and Group B used both the floating graph and the side graph.

A task starts with a 5-minute studying phase where the users explore the assigned region, followed by an evaluation quiz regarding the content just studied. All students had to perform 3 tasks, one per region. Each task had a different difficulty level, based on the number of concepts: Task 1 had 14 concepts, Task 2 had 18 concepts and Task 3 had 24 concepts, being the hardest one. All evaluation quizzes had exactly 3 questions. As for the teachers, they were not submitted to any sort of evaluation, but rather a free-hands session to explore the prototype.

5.5 Procedure

At the beginning of each session, each participant was asked to fill in an informed consent form to explain the key elements of the study and what their participation will involve, and a demographic profile form regarding their gender, education, employment (if applies), and previous VR experience, followed by a quick explanation of the structure of the session and a demonstration by the examiner on how the prototype works.

The students were first asked to train in a habituation task. Then, they were given a sequence of tasks that was randomized for each student using the Latin Squares method. All students were given 3 tasks with a maximum of 5 minutes per task, and at the end of each one, they had to complete an assessment test. The time required for finishing each assessment was measured. Figure 7 shows a student using the prototype during the user study sessions.

As for the teachers, they had free hands to use the prototype and explore deeper into the interactions and anatomic concepts presented there, for a maximum time of 10 minutes.

Figure 7: Student participating in the User Study.

5.6 Assessing Subjective Measurements

After the experimentation phase, each participant was asked to complete several questionnaires: a User Satisfaction Questionnaire (to receive feedback on the layouts and the user's preferences), a SUS questionnaire (to measure the usability of the prototype), a NASA-TLX questionnaire (to assess the task's work-load), and an IPQ questionnaire (to measure the sense of presence experienced in the virtual environment). Last but not least, the participants were submitted to a semi-structured interview regarding the use of anatomical graphs in VR, the advantages and disadvantages of using this prototype, and what changes they would suggest to improve the application. A full session lasted between 40 to 50 minutes with the students and 20 to 30 minutes with the teachers.

5.7 Statistical Analysis and Interpretation of Subjective Measurements

Statistical analysis was performed using Descriptive Statistics, the Shapiro-Wilk Test, Independent Samples t-Test, Chi-square Teste, One-Sample Wilcoxon Signed Tank Test, and Mann-Whitney U Test, all carried out using IBM SPSS Statistics 26 [2] for Windows. For all tests, a p-value of less than *alpha* = 0,05 was considered statistically significant. Since we only had 3 teachers participating, there was no need to perform any statistical analysis.

To interpret the SUS questionnaire, we computed a unique number that represents a composite measure of the overall usability of the system [6]. To calculate the NASA-TLX questionnaire, we calculated unweighted scores between 0 and 100 from a 21-item Likert scale [3] and assigned those values to a specific workload classification. [15]. To interpret the IPQ questionnaire, we divided the questions into four subscales [1], calculated the means, and assigned them certain range values [8]. Finally, the subjective data, gathered from the semi-structured interviews with the students and teachers, was analyzed through a thematic analysis method.

6 RESULTS & DISCUSSION

We divided the 30 students into two groups of 15: **Group A** was tested using only the floating graph and **Group B** was tested using both the floating graph and the side graph. Teachers also tested both layouts.

6.1 Quizzes Completion Time

The time students needed to complete the anatomical quizzes gave us some insights into the use of anatomical graphs in an immersive 3D environment. The comparison, between both groups, of the average time to complete each quiz is represented in Figure 8.

The distributions of the response times were tested for normal distribution using the Shapiro-Wilk Test. Two out of the six p-values were less than 0,05, so the assumption of normality was violated; however, the descriptive analysis showed that the characteristics of the data (*skewness* and *kurtosis* values [23]) allowed us to use parametric tests. We performed an Independent Samples t-Test with the null hypothesis (H0) defined as "The means of the quizzes completion times by using one layout or two layouts are identical". The *skewness*, *kurtosis*, means, standard deviations and the p-values from the Independent Samples t-Test are represented in Table 1.

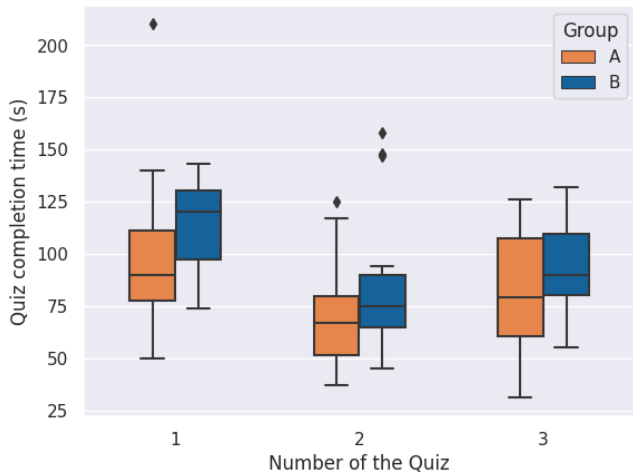

Figure 8: Comparison of quiz completion time (in seconds) for each quiz between each group.

Quiz	Group	Skewness	Kurtosis	Mean	SD	p-value
Q1	A	1,675	3,607	98,13	39,916	0,214
	B	-0,511	-0,882	113,20	22,729	
Q2	A	0,982	0,507	70,27	25,381	0,157
	B	1,201	0,246	86,67	35,582	
Q3	A	-0,021	-1,153	82,73	29,906	0,292
	B	0,066	-0,759	93,20	22,976	

Table 1: Skewness, kurtosis, mean and standard deviation (SD) of the quizzes completion time (in seconds) from each group, and p-values of the Independent Samples t-Test.

Figure 8 shows that the quiz completion time for both layouts is slightly higher than just for the conventional one. This result can be justified either by a longer time interval to mentally revive the concepts (two alternative ways to reach the same information) or the construction of mental maps to mentally organize the information. Table 1 shows that Group B took more time answering the quizzes. However, the p-values from t-Test are all above *alpha*, so we can not consider this a significant result.

6.2 Number of errors

The results of the anatomical quizzes allowed us to assess whether the use of anatomical graphs in an immersive 3D environment has the potential to be a study tool. We compared the number of incorrect answers for each question of each quiz, by each group, and we compared the percentage of incorrect answers per quiz between groups, shown in Figure 9.

These samples are independent, so we did not need to test normality: we used a non-parametric test [28], the Chi-square Test with the null hypothesis (H0) defined as "There is no relationship between the answers of both groups, for each question".

Analyzing through each question, out of 9 questions in total, Group A has more incorrect answers than Group B (5 questions versus 3 questions). Since Group B used both layouts, we inferred that anatomical graphs benefited the learning process. This goes along with Figure 9 which indicates that, overall, Group B had the best results (less percentage of incorrect answers). At last, the results from the Chi-square Test showed that only question 3.3 was statistically significant, so overall, the results do not allow us to conclude anything about the use of anatomical graphs in an immersive 3D environment as a study tool via the number of errors.

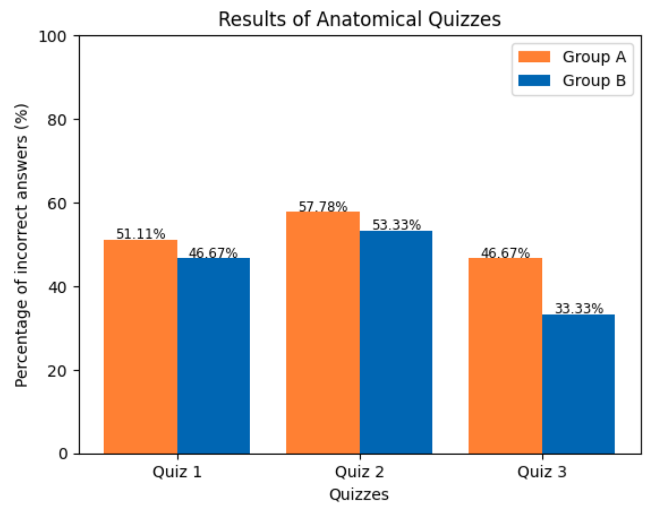

Figure 9: Comparison of the percentage of incorrect answers per quiz between groups.

6.3 Number of Concepts

The tasks follow a specific level of difficulty based on the number of concepts. From Figure 9, we can see that the percentage difference from each group between each quiz is, respectively, 4,44%, 4,45%, and 13,34%. Taking into account that the difficulty increases with the quiz (quiz 1 is the easiest and quiz 3 is the most difficult) and that the percentage differences increase as well, we can also conclude that Group B answered fewer incorrect answers and, therefore, the anatomical graphs benefited the learning process.

6.4 User Satisfaction

Participants were asked to fill in a User Satisfaction and Preference questionnaire. For the satisfaction questions, we performed the One-Sample Wilcoxon Signed Rank Test with the null hypothesis (H0) defined as "The median of each Likert item equals the hypothesized median (3,5)", and for the preference questions, we used frequency tables from descriptive statistics. The medians of the responses from all groups for the **floating graph** and their respective Wilcoxon p-values are summarized in Table 2, the medians of the responses from Group B and Teachers for the **side graph** and their respective Wilcoxon p-values are summarized in Table 3, and the frequencies from the preference questions are presented in Table 4.

Table 2 shows that all the medians were higher than 3,5, with low dispersion values (between 0 and 2) confirming that participants evaluated positively the floating graph. Table 3 proves that all medians were higher than 3,5, with low dispersion values (between

Statements	Group A	Group B	Teachers
	Mdn (IQR)	Mdn (IQR)	Mdn (IQR)
Helps locating elements anatomically.	6 (1)	6 (1)	6 (0)
Helps identifying different types of constituents.	6 (0)	5 (1)	6 (0)
Helps memorizing the constitution of the region.	5 (1)	6 (1)	6 (0)
Helps perceiving the anatomy of the region.	6 (1)	6 (1)	6 (0)
Are useful.	6 (0)	6 (1)	6 (0)
Are easy to use.	5 (1)	6 (1)	6 (0)
Help fast learning.	6 (1)	6 (1)	6 (0)
Are useful to study anatomy related to implantology.	6 (1)	6 (1)	6 (0)
Its interactivity promotes focus and learning.	6 (2)	6 (1)	6 (0)
Being able to move and rotate the 3D model is useful.	6 (0)	6 (0)	6 (0)

Table 2: Median (Mdn) and Interquartile Range (IQR) of the responses to the Likert items of the User Preference questionnaire related to the **Floating Graph**.

Statements	Group B	Teachers
	Mdn (IQR)	Mdn (IQR)
Helps locating elements anatomically.	6 (3)	6 (0)
Helps identify the different types of constituents.	6 (2)	6 (0)
Helps memorize the constitution of the region.	5 (3)	6 (0)
Helps perceive the anatomy of the region.	6 (3)	5 (0)
Is useful.	6 (2)	6 (0)
Is easy to use.	6 (1)	6 (0)
Helps fast learning.	5 (1)	6 (0)
Is useful to study anatomy related to implantology.	5 (3)	6 (0)
Its interactivity promotes focus and learning.	5 (3)	6 (0)
Interaction of both layouts helps anatomical study of the region.	6 (1)	6 (0)
Being able to grab and move the side graph is useful.	6 (1)	6 (0)

Table 3: Median (Mdn) and Interquartile Range (IQR) of the responses to the Likert items of the User Preference questionnaire related to the **Side Graph**.

Preference Question	Group B		Teachers	
	FG	SG	FG	SG
Preferred layout.	13	2	2	1
Most appealing layout.	11	4	3	0
Rank the layouts (Most Preferred)	13	2	2	1
Rank the layouts (Less Preferred)	2	13	1	2

Table 4: Answers from the Preference Questions between layouts for Group B and Teachers (FG = Floating Graph; SG = Side Graph).

0 and 3), indicating that these participants also positively evaluated the side graph. The p-values from both tests are all less than *alpha*, meaning that the results are statistically significant.

Table 4 shows that, for Group B, 13 students preferred the floating graph, 11 students found the floating graph to be the most appealing layout, and 13 students ranked the floating graph first. As for the Teachers, 2 preferred the floating graph, all 3 found the floating graph the most appealing one, and 2 ranked the floating graph first. From these results, we can conclude that both Group B and the Teachers elected the floating graph as the most useful layout.

6.5 System Usability

Participants were asked to fill in a System Usability Scale (SUS) questionnaire to measure the usability of IMPLANTIGRAPH. For the SUS score, a result above 68 is considered above average. The groups are independent, so we used a Mann-Whitney U Test with the null hypothesis (H0) defined as "The difference between the mean of the SUS score and the average score (68) is zero".

Group A's mean score was 87.50 (SD = 8.06), Group B's mean score was 87.83 (SD = 6.94), and Teachers' mean score was 90 (SD = 7.36). Since the three means are above the average, we can conclude that participants considered IMPLANTIGRAPH to have a good user interface and good usability. However, the results from the Mann-Whitney U Test showed all values of Z negative and all p-values bigger than *alpha*, so the results are not statistically significant.

6.6 Perceived Workload

For the NASA-TLX, we also performed a Mann-Whitney U Test with the null hypothesis (H0) defined as "The probability distribution of one group is the same as the probability distribution of the other group". The mean score and SD of each parameter, as well as the final score, are presented in Table 5.

NASA Parameters	Group A		Group B		Teachers	
	Mean	SD	Mean	SD	Mean	SD
MD	20,33	22,02	32,00	28,74	15,00	7,07
PD	14,00	12,14	18,00	21,74	11,67	6,24
TP	24,33	23,08	34,00	32,62	13,33	8,50
PO	16,00	22,08	15,67	24,28	28,33	33,25
EF	19,67	18,48	30,00	31,30	21,67	16,50
FR	14,33	14,81	23,33	29,19	8,33	2,36
Final Score	18,11	18,78	25,50	27,98	16,39	12,32

Table 5: Means and Standard Deviation (SD) of each NASA-TLX parameter as well as the final NASA-TLX score for each group.

Table 5 shows that for Group A the majority of the means are in the "Very Low" range, only *Mental Demand* and *Temporal Demand* enter the next value of the scale; this could be due to the amount of information to study under the 5-minute tasks. Group B has the majority in the "Low" range, with only *Physical Demand* and *Performance* in the "Very Low" range; Group B had the same information on two different layouts, so it could be overwhelming, leading to a bigger workload, hence the "Low" values; the "Very Low" values on *Phyisical Demand* and *Performance* demonstrates that using two layouts does not affect the experience. As for the Teachers, the majority are also in the "Very Low" range, with only the *Performance* and *Effort* located in the "Low" range. These results are quite different from the students. This could be due to teachers possibly having more difficulty in adapting to newer technology and having to make an extra effort to succeed.

The final scores of each group are situated in the "Very Low" (two of them) and "Low" range, which means that IMPLANTIGRAPH is not considered to have high demand levels, being an easy-to-use study and teaching tool. The results from the Mann-Whitney U test revealed all Z-values negative and all p-values bigger than *alpha*, so the results can not be considered statistically significant.

6.7 Immersive Presence

The IGroup Presence Questionnaire (IPQ) was crucial to measure the sense of being present in a virtual environment, and if the use of two layouts changes that same sense. We used a Mann-Whitney U Test with the null hypothesis (H0) defined as "The probability distribution of one group is the same as the probability distribution of the other group". The mean score and SD of each parameter, as well as the final score, are presented in Table 6.

Variables	Group A		Group B		Teachers	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
G	4,27	0,799	4,33	0,816	3,33	1,523
SP	4,27	0,704	4,27	0,961	3,80	1,474
INV	3,00	1,309	3,60	1,454	3,00	1,348
REAL	2,13	0,640	2,33	1,113	2,83	1,403

Table 6: Mean and standard deviation (SD) of the IPQ results from each group.

The G variable, for Groups A and B, is in the range of "Very Good", and for the Teachers is in the "Moderate" level, meaning that students had a higher sense of being in a virtual environment. The SP variable has the same value in both students' groups, "Very Good", but the Teachers' value decreases to "Good", indicating that the students had a major sense of being physically present in the virtual environment than the Teachers. Variable INV is in the "Moderate" range for Group A and Teachers, while Group B is on "Good". These results are a direct consequence of the location where the test sessions took place, the university clinic is a noisy environment, not allowing full concentration. Finally, the REAL environment revealed students' results in the range of "Bad" and Teachers on "Moderate", meaning that there is almost no realism in the virtual environment. In fact, there is no parallel to compare with real life: when students are studying, they do not have an interactive side graph with them to study, and when they use physical models of the jaw, these do not contain indications like the floating graph, (therefore not even existing a connection between the two), so the results, although "Bad" and "Moderate", are expected. The 3D models presented in the prototype are real bone structures, but they may not look realistic enough for students to feel that they have a real model in front of them.

The results from the Mann-Whitney U test showed all Z-values negative and all p-values bigger than *alpha*, so the results are not statistically significant.

6.8 Verbal User's Feedback

At the end of the sessions, participants were asked to answer some questions to provide detailed feedback related to IMPLANTIGRAPH and conventional methods, their advantages and disadvantages, and suggestions for further improvements.

- Complement to Conventional Studying Methods

All participants said that IMPLANTIGRAPH was a good complement to their studies. Group A stated that the information was "easy to visualize and intuitive" and the "3D perspective is a more interactive way to learn". Group B had the same opinion regarding the prototype as a whole, but as for the two layouts, the side graph was "less useful" than the floating graph. Teachers said it is "logical to have both layouts" because "one complements the other".

- Interaction and Content Benefits

Group A said that the "3D perspective" of the floating graph was "better than using books", as they could see up close the "exact locations" by "moving the model", and VR brought "spatial visualization" to a new level. Group B said that the "interaction between the two layouts" was a "good" idea, the "color scheme was a major help", and it is "good for memorization and revision". Teachers said it "eases memorization" and it is "interactive and intuitive".

- Limitations to the Prototype

Some students from both groups said that "some of the buttons, when too close together, make it hard to click", "using the prototype for an extensive period of time can be tiresome", and complained about "the controllers' sensibility". One teacher complained that "only the exterior was visible", there was no way to see "intersections or change transparency".

- Proposals for New Features

Adding the trajectories of arteries, veins, and nerves, instead of a single circular button; a feature to add a CBCT image and train a real clinical case inside the application; adding more descriptive information to the constituents; the implementation of filtering options by classification; trying to "gamify" the prototype with a topographic quiz.

- Future acceptability

When asked if participants would use the prototype, all of them said "yes": to use it while studying anatomy, or to plan real surgeries.

7 CONCLUSIONS & FUTURE WORK

7.1 Conclusions

By answering our Research Questions, we can draw conclusions for our work: the results from the number of incorrect answers and some feedback from the semi-structured questions support that anatomical VR graphs facilitate the learning process and perception of anatomical structures (RQ1); the feedback received from all participants was very positive, especially from the previous subsection *Future acceptability*, which supports that anatomical VR graphs benefit oral surgery education (RQ2); at last, we do not have data that fully supports RQ3 regarding anatomical VR graphs being a better approach than conventional methods, but subsection *Complement to Conventional Studying Methods* emphasizes that idea.

7.2 Future Work

From the feedback received as well as ideas that came up during the development, interesting features could be thought out in future work: another labeling layout, named **wrapping graph**, consisting of an extension of the conventional labeling, by adding the main structure and the different domains to the already existing constituents (all the anatomic information presented in floating tags, framing and wrapping the 3D model); a menu of filters for both layouts, having the option to select only the domains we want to study; a *collapsible* side graph, in order to choose just the subjects that are going to be studied; side graph as individual post-its.

Taking advantage of the resources that VR offers us and focusing on interactivity, we could explore the *gamification* of the application: the Tutorial Task could be transformed into an interactive follow-along tutorial; the anatomy quizzes could be performed inside the application (e.g., a task where the user must place the labels on the corresponding slots that appear around the 3D model); finally, having the option to create graphs from scratch could improve anatomy lecturing and studying by making it a more interactive and fun methodology.

REFERENCES

[1] [n. d.]. Igroup Presence Questionnaire (IPQ) overview. http://www.igroup.org/pq/ipq/index.php

[2] [n. d.]. SPSS software. https://www.ibm.com/spss

[3] [n. d.]. Task Load Index Manual. https://ntrs.nasa.gov/api/citations/20000021488/downloads/20000021488.pdf

[4] 2021. Topographic anatomy. https://www.biologyonline.com/dictionary/topographic-anatomy

[5] Sharmeen Adnan and Junhua Xiao. 2022. A scoping review on the trends of digital anatomy education. *Clinical Anatomy* (2022).

[6] Assistant Secretary for Public Affairs. 2013. System usability scale (SUS). https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

[7] Lorenzo Amabili and Nicole Sultanum. 2021. Paper Maps: Improving the Readability of Scientific Papers via Concept Maps. (2021).

[8] Muhamad Keenan Ario, Yeremia Kenny Santoso, Favian Basyari, Muhamad Fajar, Francisco Maruli Panggabean, Thomas Galih Satria, et al. 2022. Towards an implementation of immersive experience application for marketing and promotion through virtual exhibition. *Software Impacts* 14 (2022), 100439.

[9] James F Brinkley and Cornelius Rosse. 1997. The Digital Anatomist distributed framework and its applications to knowledge-based medical imaging. *Journal of the American Medical Informatics Association* 4, 3 (1997), 165–183.

[10] S Bruckner, C Turkay, and K Vrotsou. 2022. Vessel Maps: A Survey of Map-Like Visualizations of the Cardiovascular System. *STAR* 41, 3 (2022).

[11] Tony Buzan and Shengyang Guo. 2018. Mind Map Mastery: The complete guide to learning and using the most powerful thinking tool in the universe. https://www.amazon.com/Mind-Map-Mastery-Complete-Learning/dp/1786781417/ref=d_pd_sbs_sccl_3_6/139-3105081-1621957?pd_rd_w=g3tnV&content-id=amzn1.sym.3676f086-9496-4fd7-8490-77cf7f43f846&pf_rd_p=3676f086-9496-4fd7-8490-77cf7f43f846&pf_rd_r=YPSA2EDBACB5KD49YEYQ&pd_rd_wg=S2TaN&pd_rd_r=dc8aaf7f-9b66-4196-8f31-61ee26003f19&pd_rd_i=1786781417&psc=1

[12] Chien-Hsu Chen, Chun-Yen Huang, and Yin-Yu Chou. 2019. Effects of augmented reality-based multidimensional concept maps on students' learning achievement, motivation and acceptance. *Universal Access in the Information Society* 18 (2019), 257–268.

[13] Xiaojun Chen and Junlei Hu. 2018. A review of haptic simulator for oral and maxillofacial surgery based on virtual reality. *Expert Review of Medical Devices* 15, 6 (2018), 435–444.

[14] Parisa Daeijavad and Frank Maurer. 2022. Layouts of 3D Data Visualizations Small Multiples around Users in Immersive Environments. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 258–261.

[15] Anastasia Febiyani, Atik Febriani, Jauhar Ma'Sum, et al. 2021. Calculation of mental load from e-learning student with NASA TLX and SOFI method. *Jurnal Sistem dan Manajemen Industri* 5, 1 (2021), 35–42.

[16] Jakub Flotynski, Pawel Sobocinski, Sergiusz Strykowski, Dominik Strugala, Pawel Bun, Filip Górski, and Krzysztof Walczak. 2021. Semantic Representation of Domain Knowledge for Professional VR Training.. In *BIS*. 139–150.

[17] Christina Gillmann, Noeska N Smit, Eduard Gröller, Bernhard Preim, Anna Vilanova, and Thomas Wischgoll. 2021. Ten open challenges in medical visualization. *IEEE Computer Graphics and Applications* 41, 5 (2021), 7–15.

[18] Kilian Gloy, Paul Weyhe, Eric Nerenz, Maximilian Kaluschke, Verena Uslar, Gabriel Zachmann, and Dirk Weyhe. 2022. Immersive anatomy atlas: Learning factual medical knowledge in a virtual reality environment. *Anatomical Sciences Education* 15, 2 (2022), 360–368.

[19] Piotr Grad, Anna M Przeklasa-Bierowiec, Krzysztof P Malinowski, Jan Witowski, Klaudia Proniewska, and Grzegorz Tatoń. 2022. Application of HoloLens-based Augmented Reality and three-dimensional printed anatomical tooth reference models in dental education. *Anatomical Sciences Education* (2022).

[20] Uttam Grandhi and Cristián Opazo. 2021. LocalAnesthesiaVR. In *ACM SIGGRAPH 2021 Educators Forum*. 1–2.

[21] Ta-Ko Huang, Chi-Hsun Yang, Yu-Hsin Hsieh, Jen-Chyan Wang, and Chun-Cheng Hung. 2018. Augmented reality (AR) and virtual reality (VR) applied in dentistry. *The Kaohsiung journal of medical sciences* 34, 4 (2018), 243–248.

[22] Tim Joda, GO Gallucci, Daniel Wismeijer, and Nicola U Zitzmann. 2019. Augmented and virtual reality in dental medicine: A systematic review. *Computers in biology and medicine* 108 (2019), 93–100.

[23] Hae-Young Kim. 2013. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative dentistry & endodontics* 38, 1 (2013), 52–54.

[24] Martin Krajčovič, Gabriela Gabajová, Marián Matys, Patrik Grznár, L'uboslav Dulina, and Róbert Kohár. 2021. 3D interactive learning environment as a tool for knowledge transfer and retention. *Sustainability* 13, 14 (2021), 7916.

[25] Leah Labranche, Timothy D Wilson, Mark Terrell, and Randy J Kulesza. 2022. Learning in stereo: the relationship between spatial ability and 3D digital anatomy models. *Anatomical Sciences Education* 15, 2 (2022), 291–303.

[26] Zhuoshu Li, Zixin Li, Cheng Peng, Mingyi Zhao, and Qingnan He. 2022. A Bibliometric Analysis of Virtual Reality in Anatomy Teaching Between 1999 and 2022. In *Frontiers in Education*, Vol. 7. Frontiers Media SA, 874406.

[27] Meysam Siyah Mansoory, Seyyed Mohsen Azizi, Fakhrosadat Mirhosseini, Danial Yousefi, and Hedaiat Moradpoor. 2022. A study to investigate the effectiveness of the application of virtual reality technology in dental education. *BMC Medical Education* 22, 1 (2022), 457.

[28] Mary L McHugh. 2013. The chi-square test of independence. *Biochemia medica* 23, 2 (2013), 143–149.

[29] Quase médico. 2022. Uma dor de cabeça chamada 'Anatomia'? Ou nem tanto assim? https://open.spotify.com/episode/4raPgeGXmQ7lpZAy39FA8X

[30] Soraia Figueiredo Paulo, Daniel Simões Lopes, and Joaquim Jorge. 2021. 3D Reconstruction from CT Images Using Free Software Tools. In *Digital Anatomy*. Springer, 135–157.

[31] Bernhard Preim and Patrick Saalfeld. 2018. A survey of virtual human anatomy education systems. *Computers & Graphics* 71 (2018), 132–153.

[32] Rainer Schubert, Karl Heinz Höhne, Andreas Pommert, Martin Riemer, Th Schiemann, and Ulf Tiede. 1993. Spatial knowledge representation for visualization of human anatomy and function. In *Information Processing in Medical Imaging: 13th International Conference, IPMI'93 Flagstaff, Arizona, USA, June 14–18, 1993 Proceedings 13*. Springer, 168–181.

[33] Noeska N Smit, Annelot Kraima, Daniel Jansma, Marco C DeRuiter, Elmar Eisemann, and Anna Vilanova. 2016. VarVis: Visualizing Anatomical Variation in Branching Structures.. In *EuroVis (Short Papers)*. 49–53.

[34] Noeska N Smit, Kai Lawonn, Annelot Kraima, Marco C DeRuiter, Stefan Bruckner, Elmar Eisemann, and Anna Vilanova. 2019. Model-based Visualization for Medical Education and Training.. In *Eurographics (Dirk Bartz Prize)*. 5–8.

[35] Tatsuki Takano, Kazuki Takashima, Kazuyuki Fujita, Hong Guang, Kaori Ikematsu, and Yoshifumi Kitamura. 2021. A Compact and Low-cost VR Tooth Drill Training System using Mobile HMD and Stylus Smartphone. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*. 1–3.

[36] Umaiyalini Uruthiralingam and Paul M Rea. 2020. Augmented and virtual reality in anatomical education–a systematic review. *Biomedical Visualisation: Volume 6* (2020), 89–101.

[37] VictoryXR. 2021. Fisk University's 5G-Powered VR Cadaver Lab In Collaboration with VictoryXR, HTC Vive & T-Mobile. https://www.youtube.com/watch?v=VF6BhJDeWtg

[38] Zhen Wen, Wei Zeng, Luoxuan Weng, Yihan Liu, Mingliang Xu, and Wei Chen. 2022. Effects of View Layout on Situated Analytics for Multiple-View Representations in Immersive Visualization. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2022), 440–450.

[39] Ayame Yamazaki, Taku Ito, Maki Sugimoto, Soichiro Yoshida, Keiji Honda, Yoshiyuki Kawashima, Taro Fujikawa, Yasuhisa Fujii, and Takeshi Tsutsumi. 2021. Patient-specific virtual and mixed reality for immersive, experiential anatomy education and for surgical planning in temporal bone surgery. *Auris Nasus Larynx* 48, 6 (2021), 1081–1091.

[40] Eunbyul Yang, Sanghoon Park, Jeeheon Ryu, and Taehyeong Lim. 2022. How does Dental Students' expertise influence their clinical performance and Perceived Task load in a virtual Dental Lab? *Journal of Computing in Higher Education* (2022), 1–27.

[41] Lingyun Yu, Joost Ouwerling, Pjotr Svetachov, Frans h j van Hoesel, Peter MA van Ooijen, and Jiri Kosinka. 2022. VeLight: A 3D virtual reality tool for CT-based anatomy teaching and training. *Journal of Visualization* (2022), 1–14.

[42] Baoping Zhang, Sihong Li, Shuting Gao, Mingfang Hou, Hong Chen, Lulu He, Yiting Li, Yumeng Guo, Errui Wang, Rui Cao, et al. 2020. Virtual versus jaw simulation in Oral implant education: a randomized controlled trial. *BMC Medical Education* 20, 1 (2020), 1–10.

[43] Ezequiel Roberto Zorzal, Soraia Figueiredo Paulo, Pedro Rodrigues, José João Mendes, and Daniel Simões Lopes. 2021. An immersive educational tool for dental implant placement: A study on user acceptance. *International Journal of Medical Informatics* 146 (2021), 104342.