

A Large-Scale Feasibility and Ethnography Study of Screen-based AR and 3D Visualization Tools for Anatomy Education: Exploring Gender Perspectives in Learning Experience

Roghayeh Leila Barmaki^{1§}, Kangsoo Kim², Zhang Guo¹, Qile Wang¹,
Kevin Yu³, Rebecca Pearlman⁴, and Nassir Navab^{3,4}

¹ University of Delaware, Newark, DE, USA

² University of Calgary, Calgary, AB, Canada

³ Technical University of Munich, Munich, Germany

⁴ Johns Hopkins University, Baltimore, MD, USA

Abstract—Anatomy education is an indispensable part of medical training, but traditional methods face challenges like limited resources for dissection in large classes and difficulties understanding 2D anatomy in textbooks. Advanced technologies, such as 3D visualization and augmented reality (AR), are transforming anatomy learning. This paper presents two in-house solutions that use handheld tablets or screen-based AR to visualize 3D anatomy models with informative labels and in-situ visualizations of the muscle anatomy. To assess these tools, a user study of muscle anatomy education involved 236 premedical students in dyadic teams, with results showing that the tablet-based 3D visualization and screen-based AR tools led to significantly higher learning experience scores than traditional textbook. While knowledge retention didn't differ significantly, ethnographic and gender analysis showed that male students generally reported more positive learning experiences than female students. This study discusses the implications for anatomy and medical education, highlighting the potential of these innovative learning tools considering gender and team dynamics in body painting anatomy learning interventions.

Index Terms—Screen-based Augmented Reality, Collaborative Learning, Evaluation Methodologies, Human-Computer Interface, Gender and Ethnography.

I. INTRODUCTION

Human anatomy and physiology are vital parts of medical education that involve complex functional structures and movements of the human body. Comprehensive learning of anatomy and physiology provides a thorough understanding of human body function, enabling more effective treatment of abnormal or disease states [1]. However, the complexity of the course poses challenges for students in achieving their desired learning outcomes. Several factors associated with the learning experience can influence these outcomes, including the learning tools, the quality of the material, the student's prior experience, and their emotional concerns [2]–[4].

The most common practice for students in anatomy education is to use textbooks with static images, but this cannot

provide the students with a realistic first-hand and interactive experience, which may not be effective for their learning experience and performance [5]. Despite the effectiveness of traditional methods for training anatomy, such as dissection or prosection, these methods have become less feasible nowadays due to limited resources, large class sizes, and mainly the absence of face-to-face learning experiences, to name a few constraints. Due to such limitations, most medical, dental, and other allied health schools have recently declined the practical laboratory hours for anatomy [6], [7]. In anatomy and physiology education, spatial visualization is likely essential for students to learn the dynamics of anatomical structures and spatial relationships to surrounding structures. The traditional visualization of human anatomy in 2D textbook views insufficiently captures the complexity of human anatomy as students are often required to mentally reconstruct 3D spatial relationships, which presents a considerable challenge.

Virtual/augmented reality (VR/AR) can provide information on dynamics and spatial relationships interactively and intuitively by employing 3D virtual skeletons and organs and adding a virtual information layer on top of the physical body. While various medical training scenarios have used these technologies [8]–[10], the use of computer-generated 3D models, in particular, allows students to rotate and locate structures from various views and perspectives in anatomy learning. Such dynamic visualization techniques improve student visual-spatial abilities [11]–[13]. Moreover, virtual 3D visualizations offer more accessible opportunities to engage and explore anatomical structures than traditional cadaver-based learning for their repeatability and monitoring capabilities.

Body painting has been shown as an effective tool for learning anatomy and associated clinical skills [14], [15]. It is a motivating and creative experience for students that provides memorable visual images and encourages multisensory and active participation. While body painting suits all students, cultural sensitivity, gendered considerations, and careful negotiation may be necessary to ensure all students are comfortable

[§] The study was conducted during the postdoctoral fellowship of Roghayeh Leila Barmaki at Johns Hopkins University. Correspondence to rlb@udel.edu.

carrying out the activities.

Inspired by renown methods of anatomy education, in this work, we propose a user study in an actual educational laboratory setting that uses 3D visualizations in an anatomy body painting learning task. In a controlled, team-based large-scale study with 236 participants, we compare our in-house tablet-based 3D visualizations (*Tablet-3D*), and our screen-based AR (*Screen-AR*) [16] with the conventional paper-based *Textbook*. We aimed to find any interplay between the participants' performance outcomes with learning tools, gender, and group gender compositions. Our research aims to address the following research questions:

- **RQ1:** Do tablet-based visualizations and AR technology improve students' learning experience compared to the traditional textbook in anatomy education?
- **RQ2:** Do tablet-based visualizations and AR technology increase the learning outcomes, e.g., test scores or knowledge retention, in anatomy education?
- **RQ3:** Are there any particular benefits of AR technology in anatomy education experience over tablet-based visualizations?
- **RQ4:** Is the student's gender a factor to interplay with the effects of digital technology, or in general for the learning experience?

To answer these research questions, we first introduce our technological tools, tablet-based interactive visualization application and large screen-based AR tool, which can show dynamic anatomical information with interactive life-size 3D virtual models on top of a physical body. We then report the findings from our large-scale study with 236 students in teams of two who were participated in the body painting activity using three learning tools.

We found that students who used our Tablet-3D and Screen-AR conditions had more positive (anatomy) learning experiences than those who used a textbook, according to their self-reported outcomes. In addition, we analyzed the potential effects of the participant's gender on the performance, which will be elaborated more in the following.

II. RELATED WORK

A. 3D Technologies for Anatomy Education

Anatomy is a complex subject that cannot be learned only from textbooks [17]. Traditional anatomy training is based on the dissection and pro-section of the human body, which provides tangible haptic interactions and realistic environment settings [18], [19]. Despite the advantages of dissection, it equally raises concerns. Studies have shown that the learning outcomes and the quality of dissections may be affected by the quality of the material, students' prior experience, and emotional concerns [7], [20]. In particular, both inexperienced and experienced medical and healthcare students are frequently appalled by the fear of death and the unnatural smell of cadavers during the dissection [7], [21]. Instead of using deceased bodies, clay models provide an alternative solution for educators. DeHoff et al. [22] found that compared with

animal dissections, students had a better learning experience with clay models. However, clay models cannot present complex anatomical regions or the functional movements of the structures in the anatomical domain. Additionally, they impose challenges in transportation and storage. Moreover, anatomy course lab hours have gradually decreased in the past decades [6], [7], bringing more challenges to anatomy education.

As a response to the barriers and changes over time, anatomical learning platforms increasingly adapt from traditional methods to digital technology. Like any other reshaping processes, some anatomy scholars argue that dynamic visualization compensates for students' low spatial abilities by providing an explicit external representation of the system [11], [13], [23], [24]. Increasingly powerful and accessible computer hardware allow 3D visualizations to replace or supplement traditional teaching in healthcare regarding lectures, cadavers, and textbooks [25]–[28]. Donnelly et al. [29] investigated the use of Virtual Human Dissector© (VHD) software, interactive teaching tools for cross-sectional anatomy, capable of reconstructing 3D views from 2D images, in anatomy education with self-directed learning and found no significant difference when compared with a students group that learns from using prosection, models, and textbooks. Kennan & Awadh [30] discussed the effective utilization of visual 3D learning technologies as self-learning resources in the context of cross-sectional anatomy. They proposed integrating the use of the 3D VHD system with Sectra [31], a medical imaging device, to enhance the understanding of cross-sectional anatomy.

Lim et al. [32] used 3D-printed models instead of traditional cadaveric specimens during the learning of external cardiac anatomy. Although 3D printing technology can provide teaching materials, the 3D printed model is limited by the complexity of anatomical regions. Equally, researchers found that 3D visualization methods improved student performance by providing multiple anatomical views and different perspectives of 3D rotating models, even on 2D screens [25]. Mobile-based applications and web-based 3D games have also been used as efficient learning tools for the study of human skeletal, muscular, and cardiovascular systems to explore more spatial information about the 3D anatomical models [26], [27].

VR and AR techniques have been adopted into anatomy education in recent years [28], [33]–[35]. As dynamic visualizing tools, they engage students in an immersive environment with audio and visual interactions and stereoscopic 3D models to have a better functional understanding of the anatomical structure and its movement within the 3D body space [36]–[39]. Duncan-Vaidya and Stevenson [40] have found that experience from AR positively influences the learning process of skull anatomy on a similar level as traditional tools such as textbooks or plastic skull models. Increasing the frequency of learning instances and interactions with models and specimens are advantages of teaching anatomy in AR [10]. Kolla et al. [41] examined the effectiveness of VR technology in anatomy education and compared it to traditional teaching methods like lectures and cadaveric dissection. 28 first-year medical students used a VR headset to identify anatomical structures,

drew them on a virtual skeleton, and then provided feedback through surveys. Their results indicated that VR was highly supported by the students, demonstrating its potential as a valuable tool for learning human anatomy and as a useful complement. However, all study conditions were conducted within VR settings. In a different study, the AR controlled group that was randomly selected from a biochemistry course suggests that AR educational apps motivated them to understand the visualized processes [42]. Despite the popularity of using VR/AR methods in anatomy education, a review paper [43] of 152 articles did not identify conclusive evidence of their efficiency over traditional anatomy education methods.

B. Measures of Anatomy Learning Experience

To analyze the learning experience and evaluate the effectiveness of VR/AR applications, various methods for data collection and different measures were used in previous research. Kurniawan and Witjaksono [44] evaluated the usefulness of the mobile-based AR application by employing the attitude questionnaire to analyze user perception. Nainggolan et al. [17] evaluated the interactivity level of VR controller based on the user's agreement and satisfaction levels, and found that the use of the VR controller in the anatomy learning system was very interactive and satisfactory. Tanjung et al. [45] conducted a comparison learning experience study to evaluate the level of acceptability and satisfaction towards three anatomical learning systems. In our previous AR anatomy learning research [16], [33], [46], [47] and another study by Duncan-Vaidya and Stevenson [48] pre- and post-knowledge quizzes, and a usability questionnaires were used for data collection and analysis of the effectiveness of the AR tool.

C. Gender Effects in Education

In different education domains, gender differences have been discussed in recent studies. Understanding gender effects in education is not conclusive, and it varies based on different disciplines and tasks [49], [50]. In the science, technology, engineering, and mathematics (STEM) domain, some non-positive learning experiences for females have been reported due to gender biases at the technical level [51], [52]. In the business domain, females' higher management ability in group tasks was highlighted [53]–[55]. Previous research has also shown that during cognitive tests, females have better information-processing skills than males [56], [57]. In Andersson's study [58] on explicit spatial and verbal collaborative memory performance, the author reported better retention performance for females, but he argued that there was no main gender effect on team performance. Prinsen et al. [59] noted that, in computer-mediated communication [60] and computer-supported collaborative learning settings [12], [46], [61], learning performance for different genders might change by various role distributions. In this paper, we will discuss the ethnographic observations related to gender differences (from individual and group composition standpoints. We need to acknowledge that we recruited students from a large laboratory

classroom with pre-assigned teams thus, it was not feasible to form ethnographically balanced teams for all study conditions.

III. METHODS AND MATERIALS

In this section, we will describe the details of the conducted study with our proposed tablet-based 3D visualization and screen-based AR learning tools. Relevant hypotheses were established and evaluated to address our general research questions introduced in Section I. This study was approved by the Institutional Review Board (Protocol #HIRB00005021).

A. Digital Anatomy Learning Tools

To investigate the effects of tablet- and AR-based learning tools in anatomy education, we prepared in-house tablet-based 3D visualization and screen-based AR for our user study. The learning tools are presented in Figure 1. For Tablet-3D, we developed an interactive Android application for hand-held devices, which visualizes 3D virtual models of body muscles and labels on a tablet (see Figure 1(b)). The visualization of the 3D anatomy included regional and full-body anatomy models. Participants could rotate, adjust zoom levels, toggle between color-coded virtual labels for enhancing readability, and highlight relevant muscle groups for our user study (see Figure 1(c)). The developed application was installed on Samsung Galaxy Tab 2 with a 10.1-inch screen.

For Screen-AR, we deployed our screen-based AR tool onto a large TV display with a mounted Kinect v2 camera. Screen-AR used a split-screen to show an AR view on the left side and a focused view of the anatomical model on the right side. Furthermore, body tracking allowed Screen-AR to superimpose 3D life-size virtual models of human anatomical systems, such as the muscular system, and labels atop the participant's mirrored body. Participants could navigate between the relevant muscle groups using a hand-held clicker. Depending on the muscle group, the virtual camera responsible for rendering the scene on the right screen automatically follows and zooms into the selected muscles. Screen-AR consists of an Alienware 17 Laptop with GTX 1070m GPU, a 55-inch Samsung TV, and a Microsoft Kinect v2 tracking sensor, and all devices were mounted on a mobile TV cart. We predefined the same muscle groups for Tablet-3D and Screen-AR for consistency.

B. Team-based Learning Intervention

After consultation with our laboratory instructor, Dr. Pearlman about the possibilities of testing our anatomy learning tools with her large classroom, we were introduced to a mandatory team-based lab activity about human muscle painting [62]. Muscle painting, as a form of body painting, has been shown to be one of the common exercises in a medical curriculum [16], [63], and students execute the muscle painting activity typically with printed 2D visualization of human musculature from the lab manual. So, we prepared our learning tools based on these muscle painting activities, aiming to use them as a replacement for the 2D anatomy textbook and evaluate their learning efficacy and student performance. During the

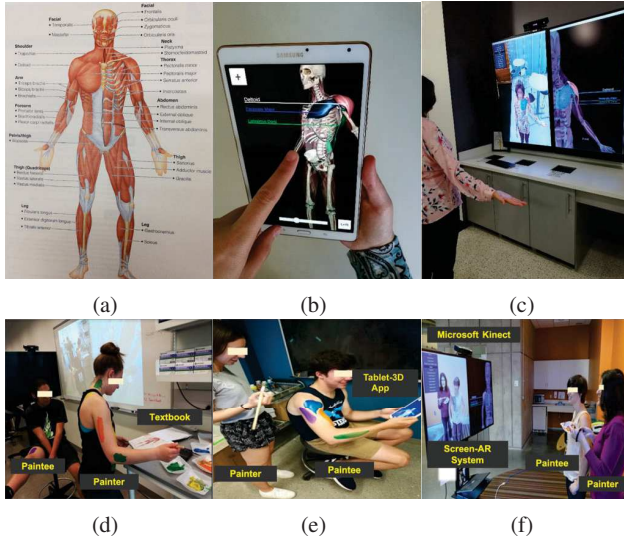


Fig. 1: Three study conditions with different learning tools: (a) Textbook, (b) Tablet-3D, and (c) Screen-AR, and the images that captured the actual anatomy learning interventions with each of those learning tools: muscle painting activity in (d) Textbook, (e) Tablet-3D, and (f) Screen-AR.

intervention, participants worked in a team of two people to collaboratively learn and teach among peers about human muscle groups through a body painting activity. They were asked to identify and paint major muscle parts on their body with washable painting supplies while using one of a randomly assigned (see Section III-C) learning tool. Once one of the participants completed the painter's role, they switched their role to be a paintee with the learning partners. We showed 40 muscles and labels to the participants to provide extensive anatomical landmarks. However, we only asked them to identify 12 major muscles (same for all conditions) and use appropriate body paint to colorize the muscles' location on their arms and legs.

C. Study Design

We conducted a user study using a between-subjects design with three *learning tool* conditions. A brief description of each *learning tool* is explained in the following.

- **Textbook:** As a traditional learning method, participants used a textbook during the team-based anatomy learning intervention described in Section III-A. They identified the corresponding muscle parts on their partner's body while checking the location of target muscles from the textbook (Figure 1(a, d)). The textbook could be carried while performing the muscle painting activity.
- **Tablet-3D:** Participants used our tablet-based 3D anatomy application during the learning intervention. The participants could carry the tablet that visualizes 3D virtual anatomy models and labels, while performing the painting activity (Figure 1(b, e)).
- **Screen-AR:** Participants used our large screen-based AR anatomy tool during the activity. They could see the virtual

TABLE I: The number of participants in each study condition.

Learning Tool	Male	Female	Total
Textbook	27	45	72
Tablet-3D	43	45	88
Screen-AR	25	51	76
Total	95	141	236

3D models and labels directly overlaid on their own bodies through the large screen (Figure 1(c, f)).

The study was performed alongside 17 anatomy lab sessions with a maximum capacity of 20 students in each lab. Considering the large scale of the study, participants could be randomly assigned to one of the three learning tool conditions by scheduling a single condition for each lab session. This assignment method is termed hierarchical or clustered randomization, which is a common practice in educational studies and clinical trials [64].

D. Participants

We initially recruited 319 student participants (male: 128, female: 191), although we needed to exclude a subset of the data (described in the following). The recruitment took place via an online flyer from a laboratory course of General Biology as part of undergraduate premedical curricula. The recruitment was part of a general biology lab course, and the students were randomly assigned by their instructor into teams of two to four to perform their assignments throughout the semester. The initial anatomy learning intervention had teams of sizes two to four, however, teams of size two were only considered for this work. Larger teams had different task distribution in team, thus excluded. So, our final subset data of interest included **236** participants (95 males and 141 females; age $M = 19.77$, $SD = 1.81$) in **118** teams of size two within three age-balanced groups with the following gender compositions. Male pairs: 62 with age $M = 19.45$, $SD = 1.035$, Female pairs: 108 with age $M = 19.71$, $SD = 1.583$, and Mixed pairs: 66 with age $M = 20.17$, $SD = 2.527$). Table I shows the number of participants per condition.

E. Procedure

The study procedure was as follows. An online flyer was sent to all undergraduate students in the General Biology lab, inviting them to participate. Students needed to complete the body painting activity regardless of our study for their lab credits, but they could opt in to participate in our study to perform it slightly differently. After the oral consent process, participants individually completed an online pre-questionnaire, which asked for their demographic information and evaluated their prior anatomy knowledge. Then they entered the learning intervention room with their pre-assigned teammates. Each team only used one of the learning tools described in Section III-C to complete the learning task, because of between-subjects study arrangements. During the painting activity, the participant in the painter's role tried to find appropriate muscles on the teammate's body (who played

the role of a paintee) to be painted while navigating different muscle parts. Participants in the Tablet-3D or textbook conditions used either the interactive application's 3D visualizations or the laboratory manual anatomy figures to complete the activity. Particularly for the AR system, the participant in the paintee's role was asked to stand closer to the Kinect body tracking sensor to have appropriate digital anatomical illustrations superimposed on their body. The painter used the muscle overlay information on the TV screen to paint. After completing the painting activity, all participants completed the online post-questionnaire individually, which asked about their interaction with the learning tools and evaluated their anatomy knowledge retention. Before concluding the study session, they presented their painted limbs to their lab assistants. The activity, including learning intervention and questionnaire completion, took approximately 20–30 minutes.

F. Measures and Hypotheses

This section describes the measures we used for the study, which were collected through questionnaires. We also introduce several hypotheses that we established based on the measures and our research questions in Section I. We used online questionnaires on the Qualtrics platform (Provo, UT) for designing and collecting pre- and post-questionnaires.

1) *Learning Experience*: We prepared nine subjective questions to examine the participants in the anatomy learning intervention using different learning tools. The questions were presented on a five-point Likert scale (1: Strongly Disagree to 5: Strongly Agree) except for the “willingness to recommend” measure, which had a scale of 0 to 10 (10 means a strong willingness). Our questions for each measure is described below. *Textbook*, *Tablet-3D*, or *Screen-AR* was replaced by *learning tool* statement, depending on the assigned condition.

- **Easy to Paint**: “Using *learning tool* was easy and straightforward for completing the muscle painting activity.”
- **Easy to Find Muscles**: “I recognized and found the location of muscles easily using *learning tool*.”
- **Satisfaction**: “Using *learning tool* for the muscle painting activity was satisfying.”
- **Learning Perception**: “I learned a lot about muscle anatomy using/interacting with *learning tool*.”
- **Enjoyment**: “I found using *learning tool* to be enjoyable.”
- **Effort to Focus**: “I had to make an effort to keep my mind on the activity.”
- **Lost Track of Time**: “Time seemed to pass very quickly during the painting activity.”
- **Willingness to Recommend**: “I would recommend that my friends use *learning tool* to study human muscles.”
- **Learning Motivation**: “Using *learning tool* increased my enthusiasm for learning more about human anatomy.”

Considering the potential benefits of interactive visualizations in our 3D visualization tool, and more intuitive and direct information display on the real body in the AR tool [65], we established the following hypotheses for the measures:

- **H1**: The participants in the Tablet-3D condition or the Screen-AR condition will have more positive ratings than

those in the Textbook condition for all the measures.

(*Textbook* < *Tablet-3D*, *Screen-AR*)

- **H2**: The participants in the Screen-AR condition will further have more positive experience than those in the Tablet-3D condition for all the measures. (*Tablet-3D* < *Screen-AR*)

As Figure 1 shows, the anatomy models in all learning tools are male-based. Compared to female students, male students can easily recognize and access the muscles according to the same anatomical structures. Also, given the prior research has shown that females tend to be more conservative about being touched [66], we established the following hypotheses to understand the role of the gender, gender composition:

- **H3**: The male participants (male pairs) in the study will report more positive ratings than the female participants (females or mixed pairs), particularly in the measures like easy to paint and easy to find muscles.

(*Female* < *Male*) and (*Mixed* <= *Females* < *Males*)

2) *Short-term Knowledge Retention*: To examine the learning performance in anatomy education, we evaluated the participant's anatomy knowledge retention. We collected the participants' anatomy knowledge scores from pre- and post-tests, and calculated the score gain by subtracting the pre-test score from the post-test score.

- **Pre-Score**: The pre-test was a matching test with five questions in a provided diagram of the human anatomy muscle system. In the anatomy diagram, 15 regions of the body were highlighted, and participants were asked to match five muscle labels provided in the pre-test with one of these 15 body regions. The number of correct matches was reported as the pre-test score in the range of [0, 5].
- **Post-Score**: The post-test had a different (lateral) view of human anatomy, with a matching test similar to the pre-test. The number of correct matches was reported as the post-test score in the range of [0, 5]. Both pre- and post-tests were designed and evaluated by anatomy instructors to be at the same level of difficulty.
- **Score Gain**: Score gain, as the difference between pre- and post-test scores, was calculated in [-5, 5] range.

Based on the positive outcome that we anticipate in the Tablet-3D and Screen-AR conditions, we established the following hypotheses similar to H1 and H2 regarding the learning performance (or the improvement of anatomy knowledge retention):

- **H4**: The participants in the Tablet-3D condition or the Screen-AR condition will have a higher score gain than those in the Textbook condition.

(*Textbook* < *Tablet-3D*, *Screen-AR*)

- **H5**: The participants in the Screen-AR condition will further have a higher score gain than those in the Tablet-3D condition. (*Tablet-3D* < *Screen-AR*)

IV. RESULTS

This section reports our analysis results considering the hypotheses we established in Section III-F. Since we are also interested in the possible effects of participant's gender on our measures, we have two factors to consider in our analysis: (1)

learning tool and (2) participant's gender. We first conducted two-way ANOVAs with these two factors to see if there is any interaction effect between the factors on both and performance measures. We did not find any significant interactions between the learning tool and the participant gender; thus, we focused on the main effects of each factor using one-way ANOVAs ($\alpha = 0.05$). Multiple comparisons with Bonferroni correction were conducted for post-hoc tests.

A. Learning Experience

Here, we report the results of our analysis on the measures for both learning tool, participant gender, and group gender composition factors. The detailed results for each measure are described below, and the overviews of the learning tool effects and the gender effects are summarized in Table II.

Easy to Paint: The one-way ANOVA for the measure of “easy to paint” showed a significant effect of the learning tool ($F(2, 233) = 7.786$, $p = 0.0005$; $\eta_p^2 = 0.063$ - medium to large effect size). The post-hoc tests revealed that Tablet-3D ($M = 4.67$, $SD = 0.656$) had a higher score than Textbook ($M = 4.26$, $SD = 0.872$; $p = 0.009$) or Screen-AR ($M = 4.18$, $SD = 1.016$; $p = 0.001$). This suggests that the participants felt the Tablet-3D condition was easier to perform the muscle painting activity than the other two conditions. The analysis for the participant gender did not show any significant effect.

Easy to Find Muscles: For the “easy to find muscles” we did not find any significant effect of the learning tool. We found that this measure was significantly different among three group gender compositions ($F(2, 233) = 5.62$, $p = 0.004$; male pairs: $M = 4.645$, $SD = 0.630$; female pairs: $M = 4.398$, $SD = 0.875$; mixed pairs: $M = 4.136$, $SD = 1.006$). Based on the post-hoc comparison with the Bonferroni test, the group with same gender compositions (males pairs and females pairs) were easier to perform the muscle painting activity than the mixed pairs ($p = 0.005$). Moreover, we found a significant effect of the participant's gender on this measure, too ($F(1, 234) = 7.065$, $p = 0.008$; $\eta_p^2 = 0.029$ - small to medium effect size). The result showed that the male participants ($M = 4.55$, $SD = 0.632$) reported a significantly higher score for the easiness of finding muscles than the female participants ($M = 4.26$, $SD = 0.937$).

Satisfaction: We did not find any significant effects of both the learning tool and the participant gender. However, this “satisfaction” measure was significantly different in gender compositions ($F(2, 233) = 3.91$, $p = 0.021$). Among three different gender compositions (male pairs: $M = 4.484$, $SD = 0.671$; female pairs: $M = 4.259$, $SD = 0.921$; mixed pairs: $M = 4.045$, $SD = 0.999$), Bonferroni test showed that male pairs had significantly higher satisfaction than mixed pairs ($p = 0.017$).

Learning Perception: For this measure, we did not find a significant effect of the learning tool, but there was a main effect of the participant gender with a statistical significance ($F(1, 234) = 6.245$, $p = 0.013$; $\eta_p^2 = 0.026$ - small to medium effect size). This showed again that the male participants ($M = 4.08$, $SD = 0.834$) had a higher score than the female

TABLE II: The summary of Learning Experience inferential results. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.)

Variable	F	p	Sig	η_p^2
Easy to Paint				
Learning tool	7.786	0.0005	***	0.063
Gender	1.854	0.175		
Easy to Find Muscles				
Gender	7.065	0.008	**	0.029
Gender composition	5.62	0.004	**	
Satisfaction				
Learning tool	0.207	0.813		
Gender	1.564	0.212		
Gender composition	3.91	0.021	*	
Learning Perception				
Learning tool	0.009	0.991		
Gender	6.245	0.013	*	0.026
Gender composition	3.63	0.028	*	
Enjoyment				
Learning tool	9.87	0.0001	***	0.078
Gender	10.008	0.002	**	0.041
Gender composition	4.03	0.019	*	
Effort to Focus				
Learning tool	1.679	0.189		
Gender	0.598	0.44		
Gender composition	1.84	0.161		
Lost Track of Time				
Learning tool	0.332	0.718		
Gender	0.413	0.521		
Gender composition	0.36	0.696		
Willingness to Recommend				
Learning tool	4.596	0.011	*	0.038
Gender	4.323	0.039	*	0.018
Gender composition	1.19	0.306		
Learning Motivation				
Learning tool	7.441	0.0007	***	0.06
Gender	4.907	0.028	*	0.021
Gender composition	2.95	0.055		

participants ($M = 3.76$, $SD = 1.062$). Also, a significant effect for the “learning perception” measure was observed ($F(2, 233) = 3.63$, $p = 0.028$) among gender compositions of male pairs ($M = 4.177$, $SD = 0.758$), female pairs ($M = 3.778$, $SD = 1.026$), and mixed pairs ($M = 3.803$, $SD = 1.084$). Bonferroni test indicated that male pairs had a significantly higher score than female pairs ($p = 0.034$).

Enjoyment: For the “enjoyment” we found main effects of the learning tool, the participant gender, and group gender composition factors. For the learning tool ($F(2, 233) = 9.870$, $p = 0.0001$; $\eta_p^2 = 0.078$ - medium to large effect size), we further compared the conditions and found that Textbook ($M = 3.78$, $SD = 1.051$) had a lower score than Tablet-3D ($M = 4.32$, $SD = 0.865$; $p = 0.001$) and Screen-AR ($M = 4.41$, $SD = 0.897$; $p = 0.001$). For the gender effect ($F(1, 234) = 10.008$, $p = 0.002$; $\eta_p^2 = 0.041$ - small to medium effect size), the result showed that the male participants ($M = 4.42$, $SD = 0.807$) had a higher score than the female participants ($M = 4.02$, $SD = 1.038$). For the gender composition (male pairs: $M = 4.468$, $SD = 0.804$; female pairs: $M = 4.037$, $SD = 1.049$; mixed pairs: $M = 4.152$, $SD = 0.932$), a significant difference was observed ($F(2, 233) = 4.03$, $p = 0.019$).

The Bonferroni post-hoc test showed that the male pairs had a significantly higher score than the female pairs ($p = 0.016$).

Effort to Focus: There were no significant effects found in the “effort to focus” measure.

Lost Track of Time: For learning tool and gender-related factors, there were also no significant effects for this measure.

Willingness to Recommend: We found a main effect of the learning tool for the “willingness to recommend” ($F(2, 233) = 4.596$, $p = 0.011$; $\eta_p^2 = 0.038$ - small to medium effect size). The post-hoc tests revealed that Textbook ($M = 6.96$, $SD = 2.185$) had a lower score than Tablet-3D ($M = 7.82$, $SD = 2.003$; $p = 0.024$) and Screen-AR ($M = 7.83$, $SD = 1.865$; $p = 0.028$). The gender effect was also found for this ($F(1, 234) = 4.323$, $p = 0.039$; $\eta_p^2 = 0.018$ - small to medium effect size), which showed that the male participants ($M = 7.89$, $SD = 1.825$) had a higher score than the female participants ($M = 7.33$, $SD = 2.164$). On the contrary, no group gender composition effect was identified.

Learning Motivation: We found the main effects of the learning tool and participant gender. The effect of the learning tool was significant ($F(2, 233) = 7.441$, $p = 0.0007$; $\eta_p^2 = 0.060$ - medium effect size), and the post-hoc tests showed that Textbook ($M = 3.57$, $SD = 1.032$) had a lower score than Tablet-3D ($M = 3.94$, $SD = 0.987$; $p = 0.044$) and Screen-AR ($M = 4.17$, $SD = 0.839$; $p = 0.001$). The gender effect was also significant ($F(1, 234) = 4.907$, $p = 0.028$; $\eta_p^2 = 0.021$ - small to medium effect size), showing that the male participants ($M = 4.07$, $SD = 0.902$) had a higher score than the female participants ($M = 3.79$, $SD = 1.020$). Based on Bonferroni test, we further compared the pairs and found that the male pairs had a significantly higher score than the female pairs for “learning motivation” ($p = 0.048$).

B. Learning Performance

To ensure valid knowledge retention evaluation, we analyzed pre-score differences among learning tool conditions and gender groups. This involved assessing the balance in prior anatomy knowledge within learning tool groups and gender compositions. We did not find any differences in the pre-score measure among the learning tool conditions. However, we found a significant difference between the male and female participants ($p = 0.019$), which showed the male participants ($M = 2.05$, $SD = 1.283$) had a higher pre-score than the female participants ($M = 1.69$, $SD = 1.070$). While controlling the pre-score as a covariate, we conducted one-way ANCOVA to investigate the effects of the learning tool, the participant’s gender, or the group gender composition; however, no significant results were found. Among different learning tool conditions, compared with the score gain ($M = 0.069$, $SD = 1.437$) and the post-score ($M = 1.92$, $SD = 1.441$) in the Textbook condition, participants in the Tablet-3D condition achieved the highest score gain ($M = 0.10$, $SD = 1.447$) with post-score ($M = 1.82$, $SD = 1.282$); Screen-AR groups achieved the second highest score gain ($M = 0.079$, $SD = 1.512$) with post-score ($M = 2.04$, $SD = 1.501$). With respect to the participant’s gender, females achieved a higher score gain ($M = 0.20$,

$SD = 1.455$) with post-score ($M = 1.89$, $SD = 1.342$) than males’ score gain ($M = -0.084$, $SD = 1.456$) with post-score ($M = 1.97$, $SD = 1.491$). Among three group gender compositions, compared with the score gain ($M = -0.145$, $SD = 1.401$) and the post-score ($M = 1.887$, $SD = 1.307$) of the male pairs, female groups achieved the highest score gain ($M = 0.269$, $SD = 1.464$) with post-score ($M = 1.852$, $SD = 1.288$); mixed pairs achieved the second highest score gain ($M = 0$, $SD = 1.488$) with post-score ($M = 2.06$, $SD = 1.654$). There are no main or interaction effects on the post-score and the score gain among the three factors: learning tool, gender, and group gender composition.

V. DISCUSSION

In this section, we summarize our findings based on the results while connecting them to our established hypotheses. We also discuss the justifications and implications of our findings and associate them with previous research.

Digital Tools Improved Learning Experience: We found several significant effects of the learning tool for some of our measures, which we reported in Section IV-A. In general, the results show that the Tablet-3D and the Screen-AR could provide more positive experience scores than the Textbook, e.g., for the measures of “enjoyment,” “willingness to recommend,” and “learning motivation,” which partly supports our **H1**. These results align with previous literature findings that showed 3D visualization technologies increased students’ engagement in anatomy learning [36], [38]. The positive scores in some of our measures could be related to the novelty of the VR/AR technology as an anatomy learning tool, and the intuitive and interactive 3D models could also be an important factor in influencing the perception of the learning experience. Some qualitative comments from the participants in the Tablet-3D condition or the Screen-AR condition post-session also support our reasoning for the positive scores. For example, some of the participants in Tablet-3D or Screen-AR said:

“Interface could be refined a little; the model was great.”

“It was an enjoyable and educational experience.”

“Super cool to be the model and see my body’s muscles.”

“It was fun to paint a muscle and helped put it into the context of reality instead of simply seeing it on a page.”

Interestingly, however, we found no significant benefits of the Screen-AR compared to the Tablet-3D, which we expected in our **H2**. Instead, we found a higher score in the Tablet-3D condition than the Screen-AR condition for the “easy to paint” measure. Based on the participants’ feedback below, we realized that some participants complained about the intermittent misalignment of virtual contents in the Screen-AR condition.

“The calibration was not great and it was hard to actually see the muscle projected onto myself and my partner.”

“Sometimes it was hard to see which muscle was being pointed.”

“It was a little hard to distinguish where in the muscle was sometimes because the overlay wasn’t actually exact/not proportional to my body.”

This implies that the participants were quite sensitive to the accuracy and reliability of body tracking for AR content registration, which could be crucial for the learning experience. Some major contributors to this AR accuracy issue include the 3D virtual models, which followed a male anatomy limb proportions rather than female, and the physical proximity of the painter and paintee in front of the body tracking system.

No Influence on Learning Performance: Positive effects of our digital tools on learning performance based on prior research was expected [16], [28], [67]. As we reported in Section IV-B, there were no significant effects of the learning tool or the participant's gender for the learning performance, e.g., the anatomy knowledge retention, which means no evidence to support our **H4** and **H5**. However, prior research showed that positive experience could increase learning performance and retention [28], [67], and our proposed 3D visualization and AR learning tools did promote more positive compared to the traditional textbook-based learning. In that sense, we are inspired to conduct further research on different learning analytic metrics. Additionally, the intervention that the participants had in our study was only a one-time session for about 10–15 minutes, which could not be enough to reveal the positive effects of AR on learning performance. As most learning science studies require longitudinal studies for their efficacy, we will propose a longitudinal study in future.

Male Users Noted More Positive Experience: Beyond investigating the learning tool effects, we also established an interesting question about gender effects in anatomy education, as introduced in **H3** in Section III-F. We found significant effects of the participant gender on various measures: “easy to find muscles,” “learning perception,” “enjoyment,” “willingness to recommend,” and “learning motivation.” For all those measures, male participants reported higher (more positive) scores than female participants, which partly supports our **H3**. Considering gender composition, similar patterns were observed on the “easy to find muscles,” “learning perception,” and “enjoyment” measures. When comparing gender compositions, we found statistically significant differences between male pairs and mixed pairs, as well as between male pairs and female pairs. Male pairs reported the highest learning experience scores. These finding also provides partial support for **H3**. We should note that this effect was not associated with a specific *learning tool*. There could be some aspects of our learning intervention that female participants did not like compared to male participants, which seemed to be the *painting part*, not the *learning tool*. Female participants in our study might be more conservative about the body painting activity than males. According to prior research [14], [66], unequal engagement of students in body painting activities may be due to cultural, social, and religious barriers concerning body image, nudity, gender, vulnerability, and embarrassment. Not surprisingly, such tendencies were reflected in the after-session feedback from some female participants.

“Painting was unnecessary, but it was cool to see the muscles on my body.” “[The activity may] take into account clothes and girls.”

“I am not a fan of body painting. However, I love the program and think that it is really cool.”

“I don’t like having paint on my skin.”

These observations emphasize the importance of gender perspectives in the design. It also recommends exploring a new learning module, one that doesn’t involve body painting.

Future Work and Limitations: The overarching goal of our study was to explore and understand the possible effects of using advanced learning technologies on anatomy learning and student performance. Identifying anatomical parts in a real body, as in our current study, is primarily acknowledged in clinical anatomy. While not immediately applicable to sessions involving cadavers or anatomy models, exploring this avenue, such as using a Kinect sensor positioned above a patient bed, presents an interesting future path. Our discussion unveils that we require multi-session, long-term, and modified studies to draw more conclusive findings for knowledge retention and task completion time. As noted, some students liked the anatomy learning tools but disliked the body painting activity, so the selection of body painting as baseline activity may have been a confounding factor. In addition, this paper covered dyadic teams and we plan to analyze our data with participants in larger teams while controlling for demographics of age, gender, and prior knowledge level to form more uniform teams. We also plan to expand student learning analysis by collecting and using multi-modal measurements for more effective and objective *learning tool* evaluations.

VI. CONCLUSION

This paper investigated the effects of digital anatomy learning tools, using our in-house tablet-based 3D visualization system and the screen-based AR application overlaying 3D anatomy structures onto the digital mirror image of students. We conducted a large-scale study with **236** premedical students, and compared our digital learning tools with conventional textbook. Our results indicated that both digital learning tools could improve the learning experience, particularly in “enjoyment”, “willingness to recommend”, and “learning motivation”. We also found that male participants generally reported more positive experiences than females. Our findings include further considerations of learners’ gender and performance in body painting anatomy education. Based on the enhanced, progressive short-term learning performance and enjoyment in our study, we expect a similar trend in the long-term retention of future 3D technology.

ACKNOWLEDGMENT

We wish to express our gratitude to the study participants and laboratory assistants. We wish to acknowledge the support from the support of the Johns Hopkins University Science of Learning Institute, and National Science Foundation for award #2321274. Any opinions, findings, and conclusions expressed in this material are those of the authors.

REFERENCES

- [1] S. Blanchard, "Anatomy and physiology," in *Introduction to Biomedical Engineering (Second Edition)*, second edition ed., ser. Biomedical Engineering, J. D. Enderle, S. M. Blanchard, and J. D. Bronzino, Eds. Boston: Academic Press, 2005, pp. 73–125.
- [2] P. J. O'Byrne, A. Patry, and J. A. Carnegie, "The development of interactive online learning tools for the study of anatomy," *Medical Teacher*, vol. 30, no. 8, pp. e260–e271, 2008.
- [3] R. A. Green, L. Y. Whitburn, A. Zacharias, G. Byrne, and D. L. Hughes, "The relationship between student engagement with online content and achievement in a blended learning anatomy course," *Anatomical sciences education*, vol. 11, no. 5, pp. 471–477, 2018.
- [4] A. Y.-C. C. Chan, O. T. Cate, E. J. Custers, M. S. van Leeuwen, and R. L. Bleys, "Approaches of anatomy teaching for seriously resource-deprived countries: A literature review," *Education for Health: Change in Learning & Practice*, vol. 32, no. 2, 2019.
- [5] B. C. Leung, M. Williams, C. Horton, and T. D. Cosker, "Modernising anatomy teaching: Which resources do students rely on?" *Journal of medical education and curricular development*, vol. 7, pp. 1–7, 2020.
- [6] K.-k. Leung, K.-S. Lu, T.-S. Huang, and B.-S. Hsieh, "Anatomy instruction in medical schools: connecting the past and the future," *Advances in Health Sciences Education*, vol. 11, no. 2, pp. 209–215, 2006, publisher: Springer.
- [7] A. Winkelmann, "Anatomical dissection as a teaching method in medical school: a review of the evidence," *Medical education*, vol. 41, no. 1, pp. 15–22, 2007, publisher: Wiley Online Library.
- [8] C. Lovis, "Mixed and Augmented Reality Tools in the Medical Anatomy Curriculum," *Digital Personalized Health and Medicine: Proceedings of MIE 2020*, vol. 270, pp. 322–326, 2020, publisher: IOS Press.
- [9] R. Marmulla, H. Hoppe, J. Mühling, and G. Eggers, "An augmented reality system for image-guided surgery: This article is derived from a previous article published in the journal International Congress Series," *International journal of oral and maxillofacial surgery*, vol. 34, no. 6, pp. 594–596, 2005, publisher: Elsevier.
- [10] M. Romand, D. Dugas, C. Gaudet-Blavignac, J. Rochat, and C. Lovis, "Mixed and augmented reality tools in the medical anatomy curriculum," *Studies in Health Technology and Informatics*, vol. 270, no. June, pp. 322–326, 2020.
- [11] T. Huk, "Who benefits from learning with 3D models? The case of spatial ability," *Journal of computer assisted learning*, vol. 22, no. 6, pp. 392–404, 2006, publisher: Wiley Online Library.
- [12] L. Lipponen, "The challenges for computer supported collaborative learning in elementary and secondary level: Finnish perspectives," 1999, publisher: International Society of the Learning Sciences (ISLS).
- [13] M. Stieff, "Mental rotation and diagrammatic reasoning in science," *Learning and instruction*, vol. 17, no. 2, pp. 219–234, 2007, publisher: Elsevier.
- [14] N. E. Cookson, J. J. Aka, and G. M. Finn, "An exploration of anatomists' views toward the use of body painting in anatomical and medical education: An international study," *Anatomical sciences education*, vol. 11, no. 2, pp. 146–154, 2018.
- [15] C. M. Diaz and T. Woolley, "'learning by doing': a mixed-methods study to identify why body painting can be a powerful approach for teaching surface anatomy to health science students," *Medical Science Educator*, vol. 31, no. 6, pp. 1875–1887, 2021.
- [16] R. Barmaki, K. Yu, R. Pearlman, R. Shingles, F. Bork, G. M. Os-good, and N. Navab, "Enhancement of Anatomical Education Using Augmented Reality: An Empirical Study of Body Painting," *Anatomical sciences education*, vol. 12, no. 6, pp. 599–609, 2019.
- [17] F. Nainggolan, B. Siregar, and F. Fahmi, "User experience in using vive controller as a controller in anatomy learning system in virtual reality environment," in *Journal of Physics: Conference Series*, vol. 1566, no. 1. IOP Publishing, 2020, p. 012096.
- [18] J. Snelling, A. Sahai, and H. Ellis, "Attitudes of medical and dental students to dissection," *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists*, vol. 16, no. 2, pp. 165–172, 2003, publisher: Wiley Online Library.
- [19] R. B. Gunderman and P. K. Wilson, "Exploring the human interior: The roles of cadaver dissection and radiologic imaging in teaching anatomy," *Academic Medicine*, vol. 80, no. 8, pp. 745–749, 2005, publisher: LWW.
- [20] R. B. Trelease, G. L. Nieder, J. Dørup, and M. S. Hansen, "Going virtual with QuickTime VR: new methods and standardized tools for interactive dynamic visualization of anatomical structures," *The Anatomical Record: An Official Publication of the American Association of Anatomists*, vol. 261, no. 2, pp. 64–77, 2000, publisher: Wiley Online Library.
- [21] J. C. McLachlan, J. Bligh, P. Bradley, and J. Searle, "Teaching anatomy without cadavers," *Medical education*, vol. 38, no. 4, pp. 418–424, 2004, publisher: Wiley Online Library.
- [22] M. E. DeHoff, K. L. Clark, and K. Meganathan, "Learning outcomes and student-perceived value of clay modeling and cat dissection in undergraduate human anatomy and physiology," *Advances in physiology education*, vol. 35, no. 1, pp. 68–75, 2011, publisher: American Physiological Society Bethesda, MD.
- [23] T. A. Hays, "Spatial abilities and the effects of computer animation on short-term and long-term comprehension," *Journal of educational computing research*, vol. 14, no. 2, pp. 139–155, 1996, publisher: SAGE Publications Sage CA: Los Angeles, CA.
- [24] J. P. Chickness, K. M. Trautman-Buckley, K. Evey, and L. Labranche, "Novel development of a 3d digital mediastinum model for anatomy education," *Translational Research in Anatomy*, vol. 26, p. 100158, 2022.
- [25] K. Yammine and C. Violato, "A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy," *Anatomical sciences education*, vol. 8, no. 6, pp. 525–538, 2015, publisher: Wiley Online Library.
- [26] N. Golenhofen, F. Heindl, C. Grab-Kroll, D. A. Messerer, T. M. Böckers, and A. Böckers, "The use of a mobile learning tool by medical students in undergraduate anatomy and its effects on assessment outcomes," *Anatomical sciences education*, 2019, publisher: Wiley Online Library.
- [27] R. R. Lemos, C. M. Rudolph, A. V. Batista, K. R. Conceição, P. F. Pereira, B. S. Bueno, P. J. Fiuza, and S. S. Mansur, "Design of a Web3D Serious Game for Human Anatomy Education: A Web3D Game for Human Anatomy Education," in *Handbook of Research on Immersive Digital Games in Educational Environments*. IGI Global, 2019, pp. 586–611.
- [28] H. Maresky, A. Oikonomou, I. Ali, N. Ditkofsky, M. Pakkal, and B. Ballyk, "Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education," *Clinical Anatomy*, vol. 32, no. 2, pp. 238–243, 2019, publisher: Wiley Online Library.
- [29] L. Donnelly, D. Patten, P. White, and G. Finn, "Virtual human dissector as a learning tool for studying cross-sectional anatomy," *Medical teacher*, vol. 31, no. 6, pp. 553–555, 2009.
- [30] I. D. Keenan and A. ben Awadh, "Integrating 3d visualisation technologies in undergraduate anatomy education," *Biomedical Visualisation: Volume 1*, pp. 39–53, 2019.
- [31] D. Barrack, D. Horn, and B. Benninger, "A step by step visual guide to using the spectra visualization table for 1st and 2nd year medical students," *The FASEB Journal*, vol. 29, pp. 692–8, 2015.
- [32] K. H. A. Lim, Z. Y. Loo, S. J. Goldie, J. W. Adams, and P. G. McMenamin, "Use of 3D printed models in medical education: a randomized control trial comparing 3D prints versus cadaveric materials for learning external cardiac anatomy," *Anatomical sciences education*, vol. 9, no. 3, pp. 213–221, 2016, publisher: Wiley Online Library.
- [33] F. Bork, R. Barmaki, U. Eck, K. Yu, C. Sandor, and N. Navab, "Empirical study of non-reversing magic mirrors for augmented reality anatomy learning," in *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2017, pp. 169–176.
- [34] J. N. Silva, M. Southworth, C. Raptis, and J. Silva, "Emerging applications of virtual reality in cardiovascular medicine," *JACC: Basic to Translational Science*, vol. 3, no. 3, pp. 420–430, 2018, number: 3 Publisher: JACC: Basic to Translational Science.
- [35] J. L. Bacca Acosta, S. M. Baldiris Navarro, R. Fabregat Gesa, S. Graf et al., "Augmented reality trends in education: a systematic review of research and applications," *Journal of Educational Technology and Society*, 2014, vol. 17, núm. 4, p. 133–149, 2014.
- [36] M. Hackett and M. Proctor, "The effect of autostereoscopic holograms on anatomical knowledge: a randomised trial," *Medical education*, vol. 52, no. 11, pp. 1147–1155, 2018, number: 11 Publisher: Wiley Online Library.
- [37] C. Jacob, S. von Mammen, T. Davison, A. Sarraf-Shirazi, V. Sarpe, A. Esmaeili, D. Phillips, I. Yazdanbod, S. Novakowski, S. Steil et al., "Lindsay virtual human: Multi-scale, agent-based, and interactive," in *Advances in Intelligent Modelling and Simulation*. Springer, 2012, pp. 327–349.

- [38] J.-M. Luursema, W. B. Verwey, P. A. Kommers, and J.-H. Annema, "The role of stereopsis in virtual anatomical learning," *Interacting with Computers*, vol. 20, no. 4-5, pp. 455-460, 2008, publisher: Oxford University Press Oxford, UK.
- [39] B. Preim and P. Saalfeld, "A survey of virtual human anatomy education systems," *Computers & Graphics*, vol. 71, pp. 132-153, 2018.
- [40] E. Duncan-Vaidya and E. Stevenson, "The Effectiveness of an Augmented Reality Head-Mounted Display in Learning Skull Anatomy at a Community College," *Anatomical Sciences Education*, 2020.
- [41] S. Kolla, M. Elgawly, J. P. Gaughan, and E. Goldman, "Medical student perception of a virtual reality training module for anatomy education," *Medical Science Educator*, vol. 30, pp. 1201-1210, 2020.
- [42] J. Barrow, C. Forker, A. Sands, D. O'Hare, and W. Hurst, "Augmented Reality for Enhancing Life Science Education," in *VISUAL 2019-The Fourth International Conference on Applications and Systems of Visual Paradigms*, 2019.
- [43] D. Chytas, M. Piagkou, T. Demesticha, G. Tsakotos, and K. Natsis, "Are extended reality technologies (erts) more effective than traditional anatomy education methods?" *Surgical and Radiologic Anatomy*, vol. 44, no. 9, pp. 1215-1218, 2022.
- [44] M. H. Kurniawan, G. Witjaksono *et al.*, "Human anatomy learning systems using augmented reality on mobile application," *Procedia Computer Science*, vol. 135, pp. 80-88, 2018.
- [45] F. Fahmi, K. Tanjung, F. Nainggolan, B. Siregar, N. Mubarakah, and M. Zarlis, "Comparison study of user experience between virtual reality controllers, leap motion controllers, and senso glove for anatomy learning systems in a virtual reality environment," in *IOP Conference Series: Materials Science and Engineering*, vol. 851, no. 1. IOP Publishing, 2020, p. 012024.
- [46] R. Barmaki and Z. Guo, "Deep neural networks for collaborative learning analytics: Evaluating team collaborations using student gaze point prediction," *Australasian Journal of Educational Technology*, vol. 36, no. 6, pp. 53-71, 2020.
- [47] F. Bork, R. Barmaki, U. Eck, P. Fallavolita, B. Fuerst, and N. Navab, "Exploring non-reversing magic mirrors for screen-based augmented reality systems," in *2017 IEEE virtual reality (VR)*. IEEE, 2017, pp. 373-374.
- [48] E. A. Duncan-Vaidya and E. L. Stevenson, "The effectiveness of an augmented reality head-mounted display in learning skull anatomy at a community college," *Anatomical Sciences Education*, vol. 14, no. 2, pp. 221-231, 2021.
- [49] L. Fernandez-Sanz and S. Misra, "Analysis of cultural and gender influences on teamwork performance for software requirements analysis in multinational environments," *IET software*, vol. 6, no. 3, pp. 167-175, 2012, publisher: IET.
- [50] J. Wegge, C. Roth, B. Neubach, K.-H. Schmidt, and R. Kanfer, "Age and gender diversity as determinants of performance and health in a public organization: the role of task complexity and group size," *Journal of Applied Psychology*, vol. 93, no. 6, p. 1301, 2008, publisher: American Psychological Association.
- [51] L. A. Meadows, D. Sekaquaptewa, and M. C. Paretti, "Interactive panel: Improving the experiences of marginalized students on engineering design teams," in *ASEE Annual Conference & Exposition: Excellence in Education*, vol. 26, 2015, p. 1.
- [52] N. Nagappan, L. Williams, M. Ferzli, E. Wiebe, K. Yang, C. Miller, and S. Balik, "Improving the CS1 experience with pair programming," *ACM SIGCSE Bulletin*, vol. 35, no. 1, pp. 359-362, 2003, publisher: ACM New York, NY, USA.
- [53] J. B. Bear and A. W. Woolley, "The role of gender in team collaboration and performance," *Interdisciplinary science reviews*, vol. 36, no. 2, pp. 146-153, 2011, publisher: Taylor & Francis.
- [54] M. De Paola, F. Gioia, and V. Scoppa, "Teamwork, Leadership and Gender," IZA Discussion Papers, Tech. Rep., 2018.
- [55] A. H. Eagly and L. L. Carli, "The female leadership advantage: An evaluation of the evidence," *The leadership quarterly*, vol. 14, no. 6, pp. 807-834, 2003, number: 6 Publisher: Elsevier.
- [56] P. Rabbitt, C. Donlan, P. Watson, L. McInnes, and N. Bent, "Unique and interactive effects of depression, age, socioeconomic advantage, and gender on cognitive performance of normal healthy older people," *Psychology and aging*, vol. 10, no. 3, p. 307, 1995, number: 3 Publisher: American Psychological Association.
- [57] K. W. Schaie and S. L. Willis, "Age difference patterns of psychometric intelligence in adulthood: generalizability within and across ability domains," *Psychology and aging*, vol. 8, no. 1, p. 44, 1993, publisher: American Psychological Association.
- [58] J. Andersson, "Net effect of memory collaboration: How is collaboration affected by factors such as friendship, gender and age?" *Scandinavian journal of Psychology*, vol. 42, no. 4, pp. 367-375, 2001, number: 4 Publisher: Wiley Online Library.
- [59] F. R. Prinsen, M. L. Volman, and J. Terwel, "Gender-related differences in computer-mediated communication and computer-supported collaborative learning," *Journal of Computer Assisted Learning*, vol. 23, no. 5, pp. 393-409, 2007, number: 5 Publisher: Wiley Online Library.
- [60] S. C. Herring, *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives*. John Benjamins Publishing, 1996, vol. 39.
- [61] E. Lehtinen, K. Hakkarainen, L. Lipponen, M. Rahikainen, and H. Muukkonen, "Computer supported collaborative learning: A review," *The JHGI Giesbers reports on education*, vol. 10, p. 1999, 1999, publisher: University of Nijmegen Nijmegen.
- [62] E. N. Marieb and P. B. Jackson, *Essentials of Human Anatomy & Physiology Laboratory Manual*. Pearson/Benjamin Cummings, 2006.
- [63] P. G. McMenamin, "Body painting as a tool in clinical anatomy teaching," *Anatomical sciences education*, vol. 1, no. 4, pp. 139-144, 2008.
- [64] L. Davis, F. G. Hamza-Lup, J. Daly, Y. Ha, S. Frolich, C. Meyer, G. Martin, J. Norfleet, K.-C. Lin, C. Imielinska, and others, "Application of augmented reality to visualizing anatomical airways," in *Helmet-and Head-Mounted Displays VII*, vol. 4711. International Society for Optics and Photonics, 2002, pp. 400-405.
- [65] T. Blum, V. Kleeberger, C. Bichlmeier, and N. Navab, "mirracle: An augmented reality magic mirror system for anatomy education," in *2012 IEEE Virtual Reality Workshops (VRW)*. IEEE, 2012, pp. 115-116.
- [66] G. M. Finn and J. C. McLachlan, "A qualitative study of student responses to body painting," *Anatomical sciences education*, vol. 3, no. 1, pp. 33-38, 2010.
- [67] D. T. Nicholson, C. Chalk, W. R. J. Funnell, and S. J. Daniel, "Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model," *Medical education*, vol. 40, no. 11, pp. 1081-1087, 2006, publisher: Wiley Online Library.