3D Learning in anatomical and surgical education in relation to visual-spatial abilities

Bogomolova, K.

3

The effect of stereoscopic augmented reality visualization on learning anatomy and the modifying effect of visual-spatial abilities: a double-center randomized controlled trial

Katerina Bogomolova, Ineke J.M. van der Ham, Mary E.W. Dankbaar, Walter W. van den Broek, Steven E.R. Hovius, Jos A. van der Hage, Beerend P. Hierck

Anatomical Sciences Education 2020; 13:558-567

ABSTRACT

Background

Monoscopically projected three-dimensional (3D) visualization technology may have significant disadvantages for students with lower visual-spatial abilities (VSA) despite its overall effectiveness in teaching anatomy. Previous research suggests that stereopsis may facilitate a better comprehension of anatomical knowledge. This study evaluated the educational effectiveness of stereoscopic Augmented Reality (AR) visualization and the modifying effect of VSA on learning.

Methods

In a double-center randomized controlled trial, first and second-year (bio)medical undergraduates studied lower limb anatomy with a stereoscopic 3D AR model (n = 20), a monoscopic 3D desktop model (n = 20) or two-dimensional (2D) anatomical atlas (n = 18). VSA were tested with Mental Rotation Test (MRT), Paper Folding Test and Mechanical Reasoning Test. Anatomical knowledge was assessed by the validated 30-item paper post-test.

Results

The overall post-test scores in the stereoscopic 3D AR group (47.8 %) were similar to those in the monoscopic 3D desktop group (38.5 %; p = .081) and the 2D anatomical atlas group (50.9 %; p = 1.00). When stratified by VSA test scores, students with lower MRT scores achieved higher post-test scores in the stereoscopic 3D AR group (49.2 %) as compared to the monoscopic 3D desktop group (33.4 %; p = .015) and similar to the scores in the 2D group (46.4 %; p = .99). Participants with higher MRT scores performed equally well in all conditions.

Conclusions

It is instrumental to consider an aptitude-treatment interaction caused by VSA when designing research into 3D learning. Further research is needed to identify contributing features and the most effective way of introducing this technology into current educational programs.

INTRODUCTION

Anatomical knowledge among undergraduate medical students and recently graduated doctors has repeatedly been reported to be insufficient.[1-5] One of the main reasons is the decrease in anatomy teaching time in undergraduate education, related to the increasing costs and limited availability of cadavers, and the time pressure on the curriculum have led to a decreased exposure to traditional cadaveric dissections.[4,6-9] Although, the educational value is being debated[7], cadaveric dissections provide a complete visual and tactile learning experience of anatomy which is three-dimensional (3D) by nature. Features such as stereopsis (visual sense of depth), dynamic exploration (the possibility to view the object of study from different angles), and haptic feedback (sense of touch) are crucial for the engagement in 3D anatomy.[8,9]

In search of additional educational resources, computer assisted resources have been widely explored in anatomical education. A considerable number of studies have evaluated the effectiveness of digital 3D anatomical models which can be explored on a two-dimensional (2D) screen of a regular computer, smartphone, or tablet. In an extended meta-analysis of these studies, Yammine and Violato[10] concluded that three-dimensional visualization technology (3DVT) is effective in improving factual (effect size of 0.50) and spatial (effect size of 0.30) anatomical knowledge. However, despite of the overall positive effect on learning, 3DVT appears to have significant disadvantages for students with low visual-spatial abilities (VSA).[11-15] The disadvantages are well known in the research field of 3D learning and were first described by Garg and colleagues. [11-13] In these studies, VSA significantly affected the learning process of spatial anatomy showing a great disadvantage for low performing students. Viewing an unfamiliar 3D object from multiple angles would be challenging for these students due to evidence that 3D objects are remembered as key view based 2D images.[13,14,16,17]

However, when traditional digital 3D models are viewed stereoscopically by projecting a slightly shifted image to the left and right eye, the disadvantages for low VSA students seem to disappear. Cui and colleagues have evaluated the effectiveness of a stereoscopic 3D view of the head and neck vascular anatomy in comparison to 2D representations of the same anatomical model.[18] They reported a better performance of undergraduate medical students after learning anatomy with a stereoscopic 3D model. Most importantly, students with low VSA have improved their knowledge test scores to a level comparable to that demonstrated by the high VSA students. The role of stereopsis has also been evaluated by Luursema and colleagues[19-21] within various 3D environments, such as virtual reality and stereoscopic projection on a computer with the use of 3D shutter glasses. Although the stereoscopic view of an anatomical model has had a positive effect only on one of the two post-tasks, the interaction between VSA and the stereoscopic condition remained significant.[20] Overall, stereoscopic 3DVT appears to have a positive effect on learning as recently demonstrated by Hackett and Proctor.[22] Their intervention concerned an autostereoscopic holographic visualization of a cardiac 3D model which has been compared to a monoscopic desktop view and 2D printed images of the model. Students in the intervention group scored significantly higher on the anatomical knowledge test and have reported a significantly lower cognitive load in comparison to both control groups. However, a possible interaction between intervention and VSA has not been evaluated. The positive role of stereopsis has also been shown when a physical model of the pelvis was compared to a monoscopic 3D model by Wainman and colleagues.[23] Authors have concluded that stereopsis, and not haptic feedback, primarily contributed to the improved knowledge scores when learning with a physical model.

In regard to these findings, two aspects come into play. Firstly, beneficial effects of stereopsis support the evidence that 3D mental representations depend on the nature of the input by activating different regions of the brain, and might contain spatial information instead of key-view based 2D images alone.[19,24-26] Stereopsis might therefore facilitate a better comprehension of anatomy especially among students with lower levels of VSA. Secondly, the reported differences in learning effect between students with lower and higher levels of VSA in various interventions possibly reflect an aptitude–treatment interaction. An aptitude-treatment interaction occurs when a student's attribute, e.g., visual-spatial abilities, predicts different outcomes for different treatments.[27] Such interaction is only detectible when the outcomes are stratified by the variable or when the variable is included in the regression analysis as an interaction term (variable x intervention), as demonstrated by Luursema and colleagues and Cui and colleagues.[18,20]

Augmented reality in anatomical education

Augmented reality (AR) is a new generation of 3DVT technology that is eagerly being explored in the field of anatomical education and research in recent years.[27,28] It gained popularity due to its ability to combine 3D computer-generated virtual objects with physical environment. This enables learners to interact with each other and with the digital environment using mobile devices, such as smart phones and tablets, or, more recently, head-mounted displays (HMD) such as AR and virtual reality (VR) devices. Whether the anatomy can be perceived in a real three-dimensional plane, depends on the type of device. Visualization of 3D content from flat screens is usually obtained monoscopically with various interactive features added to the digital overlay provided by these devices.[29-31] HMD can provide an interactive and stereoscopic way of 3D visualization (*Supplementary material 2*). With AR technology, such as with the Hololens®, the most distinguishing feature is the ability to perceive an anatomical model in a real three-dimensional plane without losing sense of the user's own environment. Dynamic exploration, an object centered view, enables users to walk around the stereoscopic model and explore it from all possible angles. The use of this technology has been reported in the surgical field of preoperative planning and tumor localization.[32] The educational effectiveness of this technology for teaching anatomy has not been evaluated yet. For the purpose of this study an AR application *DynamicAnatomy* was developed at the department of Anatomy and Embryology at Leiden University Medical Center and the Centre for Innovation of Leiden University. This application provides a dynamic stereoscopic 3D view on the lower limb including the musculoskeletal anatomy. Further specification of the application is provided in the Methods section.

Objectives and aims

The aim of this study was to evaluate the learning effect of an anatomical stereoscopic 3D AR model of the lower leg among medical undergraduates when compared to a monoscopic 3D desktop model and 2D anatomical atlas. The secondary objective was to evaluate whether VSA would modify the observed learning effect. Additionally, the study aimed to evaluate the student's experience of learning anatomy in AR. The authors hypothesized that the stereoscopic 3D AR model is more effective in improving anatomical knowledge than the monoscopic 3D desktop model and the 2D anatomical atlas, and that students with lower levels of VSA benefit most from the stereoscopic 3D view of the model.

MATERIALS AND METHODS

Study design

A double-center randomized controlled trial was conducted at the Leiden University Medical Center (LUMC) and the Erasmus University Medical Center Rotterdam (EMC), the Netherlands in the spring of 2018 (Figure 1). The study protocol was approved by the Institutional Review Board at the Leiden University (registration no. CEP17-1215/420). Participation was voluntary and written consent was obtained from all participants.

Study population

Participants were a volunteer sample of first- and second-year undergraduate students of Medicine and Biomedical Sciences at the LUMC and EMC and were recruited through flyers and announcements during the lectures. The study took place prior to the anatomy courses on the musculoskeletal system of the limbs, ensuring limited knowledge of the lower limb anatomy among all participants. Students who had already taken part in this course were excluded. The baseline knowledge was not assessed to avoid extra burden for students and possible influence on learning during the intervention and the performance on the post-test.[32] Participation in the study did not interfere with the curriculum and the assessment results did not affect student's academic grades. Participants received a compensation of fifteen euros at the completion of the experimental session.

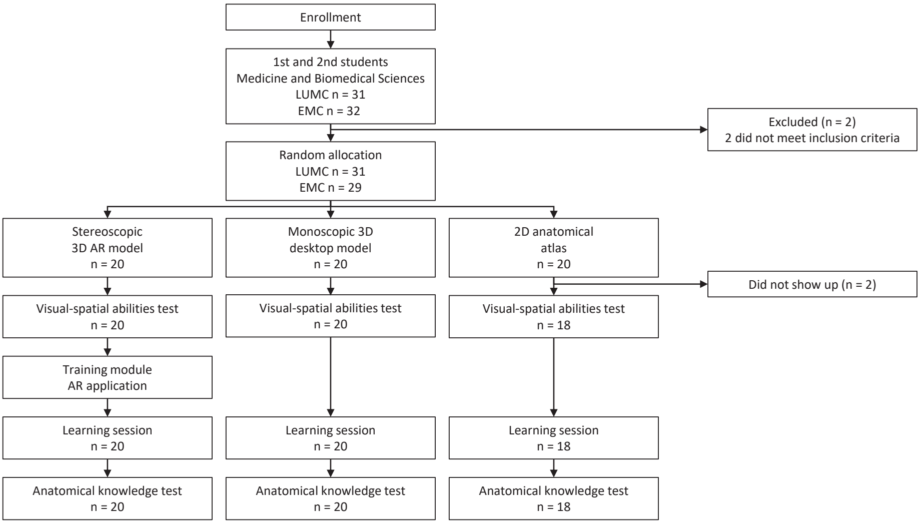

Figure 1. Flowchart of study design. LUMC, Leiden University Medical Center; EMC, Erasmus Medical Center Rotterdam; n, number of participants; AR, Augmented Reality; 3D, three-dimensional; 2D, two-dimensional.

Randomization

Participants who consented to participate were randomly allocated to either the (1) stereoscopic 3D AR model group, (2) monoscopic 3D desktop model group or (3) 2D anatomical atlas group. Students were assigned an identification number, and these were randomly allocated to the three groups using an Excel Random Group Generator. Blinding of participants was impossible since the intervention was apparent to the students.

Educational interventions

For the purpose of this study an AR application *DynamicAnatomy* (https://www.microsoft.com/en-us/p/dynamicanatomy/9nwlj4qq053p?SilentAuth=1&activetab=pivot:overviewtab) for Microsoft Hololens® (Version 1, Microsoft, Redmond, Washington, USA) was developed at the Department of Anatomy and Embryology at Leiden University Medical Center and the Centre for Innovation of Leiden University (www.mr4education.com). The application represented a dynamic and fully interactive stereoscopic 3D model of the lower leg. The model was presented as a three-dimensional virtual object in the physical space (*Supplementary material 2*). The Hololens glasses are transparent which enabled participants to stereoscopically interact with the model without losing sense of their own physical environment. A unique feature included an object centered view, i.e., dynamic exploration, which enabled participants to walk around the 3D model and explore it from all possible angles. Participants navigated through the user interface and selected desirable functions by making specific hand gestures or giving a voice command. Active interaction included size adjustments, showing or hiding structures by group or individually, visual and auditory feedback on structures and anatomical layers, and animation of the ankle movements.

With the gaze function switched on, the text of the anatomical descriptions appeared next to the highlighted structure. The anatomical layers included musculoskeletal, connective tissue, and neuro-vascular systems. During this experiment, study participants focused on the musculoskeletal system. Prior to the experiment, participants completed a 10-minutes training module, without anatomical content, to get familiar with the use of the application and device.

For the intended comparison, a Windows desktop application was developed with all the features of *DynamicAnatomy*. The desktop application included the identical anatomical model of the lower limb which was now displayed monoscopically on a 2D computer screen. The model could be rotated along the Y axis in both directions with a slide-bar using a computer mouse *(Supplementary material 3)*. All other features such as voice control, auditory feedback, and scaling were unchanged (Table 1).

Table 1. An overview of available features of the three educational interventions.

Feature	Stereoscopic 3D AR model	Monoscopic 3D desktop model	2D anatomical atlas
Stereopsis	+	-	-
Dynamic exploration	+	-	-
Active user interaction	+	+	-
Animation of the ankle	+	+	-

+ , a feature is present; -, a feature is not present; AR, augmented reality; 3D, three-dimensional; 2D, two-dimensional.

In the 2D anatomical atlas group, study material included selected handouts from an anatomy atlas[33] and an anatomy textbook[34] covering anatomy of the musculoskeletal system. The selection consisted primarily of 2D images of bones and muscles of the lower leg and ankle movements with short descriptions. Each handout included an index for the ease of navigation. In all groups the anatomical descriptions were limited to the names of the structures. No additional textual descriptions were provided.

Learning objectives and instructional activities

Participants received a handout with a description of the learning goals (identical for each group) and instructions for the learning session (specific for their group). Both were developed based on the constructive alignment theory to ensure the alignment between the intended learning outcomes, instructional activities and knowledge assessment *(Supplementary material 4A, 4B).*[35] The learning goals were formulated and organized according to Bloom's Taxonomy of Learning Objectives.[36] An independent expert outside of the anatomy verified the alignment between the learning goals and the assessment according to the constructive alignment theory and Bloom's Taxonomy of Learning Objectives. Learning goals included memorization of the names of bones and muscles (factual knowledge), understanding the function of the muscles based on their origin and insertion (functional knowledge), and location and organization of these structures in relation to each other (spatial knowledge). Students were free to follow the provided instructions or to choose their own way of achieving the learning goals. Duration of the learning session was 45 minutes.

Assessment of VSA

VSA were assessed prior to the start of the learning session. Mental visualization and rotation, as the main components of VSA, were assessed by the 24-item Mental Rotation Test (MRT), previously validated by Vandenberg and Kuse[37] (1978) and redrawn by Peters and colleagues[38] (*Supplementary material 5A*). This psychometric test is being widely used in the assessment of VSA and has repeatedly shown its positive association with anatomy learning and assessment.[39,40] The post-hoc level of internal consistency (Cronbach's alpha) of the MRT in this study was 0.88. Mental visualization and transformation, as other components of VSA, were measured by the 10-item Paper Folding Test (PFT), previously validated by Ekstrom and colleagues[41] *(Supplementary material 5B).* The post-hoc level of internal consistency (Cronbach's alpha) of the PFT in this study was 0.76. Additionally, mechanical reasoning was measured by a standardized 12-item Mechanical Reasoning Test (MR), developed for this experiment at the Department of Neuropsychology *(Supplementary material 5C).* The post-hoc level of internal consistency (Cronbach's alpha) of the MR test in this study was 0.76. The duration of the assessment was three minutes for each test. After three minutes all students were instructed to collectively move on to the next test even if they did not finish all the items.

Anatomy knowledge assessment

The learning effect was evaluated by a 30-item knowledge test. The test consisted of a combination of twenty extended matching questions and ten open-ended questions. The knowledge was assessed in the factual (i.e., memorization/identification of the names of bones and muscles), functional (i.e., understanding the function of the muscles based on their course, origin and insertion) and spatial (i.e., location and organization of structures in relation to each other) knowledge domains (*Supplementary material 6*). Content validation was performed by two experts in the field of anatomy and plastic and reconstructive surgery. The test was then piloted among twelve medical students for item clarity. The post-hoc calculated level of internal consistency (Cronbach's alpha) was 0.78. The duration of the assessment was 30 minutes.

Evaluation of learning experience

Participants' learning experience was evaluated by a standardized self-reported questionnaire. The evaluation included items on study time, perceived representativeness of the test questions, perceived knowledge gain, usability of and satisfaction with the provided study materials. Response options ranged from "very dissatisfied" (1 point) to "very satisfied" (5 points) on a five-point Likert scale.

Statistical analysis

Participant's baseline characteristics were summarized using descriptive statistics. The differences in baseline measurements were assessed with a one-way ANOVA for differences in means and X^2 test for differences in proportions. The normal distribution was assessed with Shapiro Wilk Test of Normality in combination with the Normal Q-Q Plots. The differences in mean percentages of correct answers on the anatomy knowledge test between groups were assessed with one-way ANOVA including mean percentages of correct answers as a dependent variable and intervention group as a fixed factor. In case of a significant difference, a post-hoc Bonferroni test was performed to identify the pairs of means that differ. The obtained p values were adjusted for multiple comparisons with a Bonferroni correction (*P* value*k). The results were stratified by MRT, PFT and MR test scores to evaluate possible aptitude-treatment interaction between VSA and type of intervention. Additionally, a ANCOVA was performed to evaluate the interaction in a linear regression analysis. Anatomy knowledge test score was included as a dependent variable, intervention group as a fixed factor, VSA test score as a covariate, and 'VSA test score' x 'intervention group' as in interaction term. The effect size (Cohen's d) of the differences in anatomy knowledge test scores between groups was calculated using the mean scores and standard deviations of two groups.[42] All analyses were performed using SPSS statistical software package version 23.0 for Windows (IBM Corp., Armonk, NY). Statistical significance was determined at the level of $p < .05$.

RESULTS

A total of sixty participants were included in the study. Two participants allocated to the 2D anatomical atlas group did not show up for the experiment. The 2D anatomical atlas group, therefore, consisted of 18 participants. Participants were not aware of their allocation to one of the three groups in advance but were informed prior to the start of the experiment. Table 2 shows the baseline characteristics of the 58 participants.

Table 2. Baseline characteristics of the included participants.

	Stereoscopic 3D AR model	Monoscopic 3D desktop model	2D anatomical atlas	p value
	n = 20	n = 20	n = 18	
Sex, n (%)				
Male	8 (40.0)	6 (35.0)	7 (39.0)	.773
Female	12 (60.0)	13 (65.0)	11 (61.0)	
Age, mean ± SD, y	18.5 ± 0.8	18.7 ± 1.0	18.7 ± 0.8	.720
Medical center, n (%)				
Leiden University MC	10 (50.0)	11 (55.0)	10 (55.6)	.929
Erasmus University MC Rotterdam	10 (50.0)	9 (45.0)	8 (44.4)	
Study, n (%)				
Medicine	17 (85.0)	16 (80.0)	14 (77.8)	.842
Biomedical sciences	3 (15.0)	4 (20.0)	4 (22.2)	
Study year, n (%)				
1st year	17 (85.0)	18 (90.0)	16 (88.9)	.879
2nd year	3 (15.0)	2 (10.0)	2 (11.1)	
Visual-spatial abilities score, mean ± SD				
Mental Rotation Test	7.1 ± 2.9	6.0 ± 2.4	8.4 ± 2 .1	.090
Paper Folding Test	6.2 ± 1.8	6.5 ± 2.6	7.6 ± 2.2	.104
Mechanical Reasoning Test	9.3 ± 2.6	9.3 ± 2.1	9.3 ± 3.2	.990

$p < .05$ with a Bonferroni correction for multiple comparison is considered significant. Minimal and maximal scores range between 0-24 for the Mental Rotation Test, 0-10 for the Paper Folding Test and 0-12 for the Mechanical Reasoning Test. n, number of participants; AR, Augmented Reality; 3D, three-dimensional; 2D, two-dimensional; SD, standard deviation; y, years; MC, medical center.

Overall scores on anatomy knowledge assessment

The scores are presented as mean percentages of correct answers. As shown in Figure 2, the stereoscopic 3D AR group (47.8 %, SD ± 9.8) performed equally well on the knowledge test as the monoscopic 3D desktop group (38.5 %, SD ± 14.3; $F(2,54) = 4.79$; $p = .081$) and the 2D anatomical atlas group (50.9 %, SD ± 13.8; $F(2,54) = 4.79$; $p = 1.00$). The 2D anatomical atlas group, however, outperformed the monoscopic 3D desktop group ($F(2,54) = 4.79$; $p = .042$).

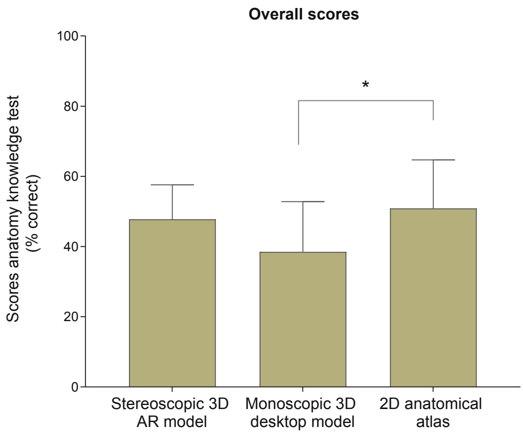

Figure 2. Differences in overall mean percentages correct answers on the anatomy knowledge test between three educational interventions. a, $p < .05$ analysis of variance with a Bonferroni correction for multiple comparison. MRT, Mental Rotation Test; AR, augmented reality; 3D, three-dimensional; 2D, two-dimensional.

Stratified by VSA scores

When total scores on the anatomy knowledge test were stratified by Mental Rotation Test (MRT), Paper Folding Test (PFT) and Mechanical Reasoning (MR) test scores, only the MRT scores did significantly impact the outcomes in all three conditions. Students who scored below the mean were assigned to the MRT-low group (n=31) and students who scored above the mean were assigned to the MRT-high group (n=26). As shown in Figure 3, the MRT-high group performed equally well in each of the three intervention groups ($F(2,23) = 0.83$, $p = .448$). However, among MRT-low participants significant differences were found between groups. The stereoscopic 3D AR group (49.2 %, SD ± 9.5) significantly outperformed the monoscopic 3D desktop group (33.4 %, SD ± 11.5; $F(2,28) = 6.59$, $p = .015$, Cohen's $d = 1.54$), and performed equally well as the 2D anatomical atlas group (46.4 %, SD ± 14.5; $F(2,28) = 6.59$, $p = .990$, Cohen's $d = 0.24$). Although, students achieved higher scores in the 2D anatomical atlas group than in the monoscopic 3D desktop group with a moderate effect size (Cohen's $d = 1.00$), the observed difference was not significant (p

= .080). The MRT-low group performed significantly worse than the MRT-high students in the monoscopic 3D desktop group (33.4 %, SD ± 11.5 vs. 49.7 %, SD ± 13.9; p = .015, Cohen's *d* = -1.3) However, they performed equally well in the stereoscopic 3D AR and 2D anatomical atlas groups.

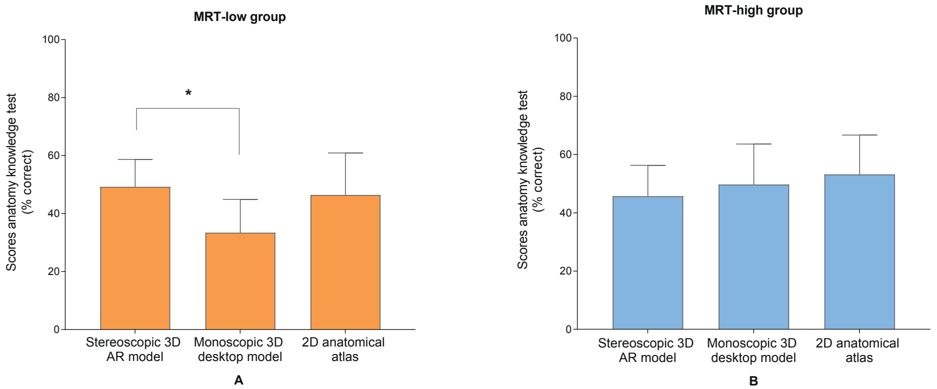

Figure 3. Differences in overall mean percentages correct answers on the anatomy knowledge test between three educational interventions stratified by Mental Rotation Test scores. A, Students who scored below the mean were assigned to the MRT-low group (n = 31) and B, students who scored above the mean were assigned to the MRT-high group (n = 26). [a] p < .05 analysis of variance with a Bonferroni correction for multiple comparison. MRT, Mental Rotation Test; AR, augmented reality; 3D, three-dimensional; 2D, two-dimensional.

The observed differences strongly indicate an aptitude-treatment effect caused by VSA. This phenomenon occurs when the effect of an intervention is different in groups of subjects with different characteristics. Therefore, the observed interaction between the MRT scores and the intervention groups was additionally checked in a linear regression analysis. The interaction term 'MRT score' x 'intervention group' showed a marginal trend towards significance (F(2) = 3.04; p = .05). Including PFT and MR test scores as a covariate and an interaction term did not have any significant impact on the outcomes.

Evaluation of learning experience

As shown in Table 3, participants in the stereoscopic 3D AR group enjoyed the learning session more than the participants in other two groups (4.8 ± 0.4 vs. 3.4 ± 0.8 vs. 2.4 ± 0.9; F(2,54) = 50.3, p = .003). Participants found the application easy and intuitive to use and would recommend it to their fellow students. In all three groups participants reported that their knowledge about anatomy of the lower leg was improved (4.3 ± 0.6 vs. 4.1 ± 0.9 vs. 4.1 ± 0.8; F(2,54) = 0.6, p = .574).

Table 3. Differences in students' learning experience between three educational interventions.

	Stereoscopic 3D AR group	Monoscopic 3D desktop model	2D anatomical atlas	p value
	n = 20	n = 20	n = 18	
The study time was long enough to study the required number of anatomical structures	3.0 ± 0.9	2.3 ± 0.7	2.6 ± 0.9	.192
The questions in anatomy test were representative for the studied material	3.8 ± 0.6	3.8 ± 0.6	3.7 ± 0.7	.709
I enjoyed studying with …	4.8 ± 0.4§	3.4 ± 0.8§	2.4 ± 0.9§	.003*
Learning material was easy to use	4.3 ± 0.6ǂ	3.4 ± 0.9ǂ	3.0 ± 0.8ǂ	.009*
My knowledge about anatomy of the lower leg is improved after studying with …	4.3 ± 0.6	4.1 ± 0.9	4.1 ± 0.8	.574
I would recommend studying with … to my fellow students	4.6 ± 0.5§	3.7 ± 0.8§	2.4 ± 0.9§	.003*

Response options on a 5-point Likert scale ranged from 1 = very dissatisfied to 5 = very satisfied on a 5-point Likert scale. Average scores are expressed in means (± SD). * $p < .05$ analysis of variance with a Bonferroni correction for multiple comparison; § significant difference between all the three groups. ǂ significant difference between (1) Stereoscopic 3D AR model and monoscopic 3D desktop model group; (2) Stereoscopic 3D AR model and 2D anatomical atlas group. n, number of participants; AR, Augmented Reality; 3D, three-dimensional; 2D, two-dimensional; SD, standard deviation.

DISCUSSION

This study aimed to investigate the educational effectiveness of learning with stereoscopic AR visualization technology and to evaluate whether VSA would modify the learning effect.

Firstly, the observed aptitude-treatment interaction caused by VSA needs to be addressed in more depth. The results showed significant differences in learning effect upon interventions using 2D and 3D learning materials among participants with lower and higher VSA scores as measured by the MRT. These differences were detectible only after stratification of the overall results pointing towards an aptitude-treatment interaction, also referred to as 'effect measure modification'.[27, 43,44] This phenomenon occurs when the effect of an intervention is different in groups of subjects with different characteristics and is different from the effect of a confounder. In current analyses, when VSA were treated only as a confounder, in the absence of stratification, the differences between monoscopic and stereoscopic conditions for different levels of VSA were not evident. This means that an adjustment for this confounder by the study design (e.g., randomization) or statistical analysis (e.g., including it only as a co-variate in the regression analysis), will still not be sufficient, and the results can still be misleading.

Secondly, the monoscopic 3D desktop model group only showed a lower learning effect in the MRT-low group. These findings are supported by previous research in the effectiveness of monoscopic 3D visualization technologies with disadvantages for students with low VSA.[11-15] It has been hypothesized that three-dimensional objects are memorized as key view based two-dimensional images.[13,45] Viewing an unfamiliar 3D object from multiple angles, could therefore lead to an increase in extraneous cognitive load.[16,17,46] The beneficial effect of stereoscopic visualization of a 3D object could be explained by the fact that mental representations depend on the nature of the input.[20,24, 25] In that case, mental representations do not primarily consist of key view based 2D images, but they might also include spatial information. This is further supported by the evidence that disparity processing occurs in different visual pathways of the human brain.[26] This means, that while a monoscopic 3D desktop view and 2D anatomical atlas images would stimulate key view based 2D mental images, a stereoscopic 3D model would stimulate structural 3D mental representations. Stereopsis might then avoid the increase in extraneous cognitive load and therefore facilitate a better comprehension of 3D anatomy in students with lower levels of VSA.

As dynamic exploration was the second distinguishing feature of the stereoscopic 3D AR model, it may also have contributed to the positive learning effect. Being able to walk around the model with its own reference point can create an additional sense of depth. On the other hand, the object centered view is different from the egocentric view where the user moves the objects in their field with virtual tools, as was the case in the monoscopic 3D desktop group. The egocentric control can affect visual-spatial skills where the hands are involved in imagining the rotation of objects. Future research is needed to evaluate how these different types of view in a 3D environment affect spatial processing during learning. This should be performed in an identical environment using the same medium, configuration and presentation.[27] This eliminates all possible confounding effects of additional features such as hand gestures, that can vary between different types of media.

Thirdly, participants in the 2D anatomical atlas group achieved anatomy knowledge test scores similar to those in the stereoscopic 3D AR model group. This unexpected effect can be hypothetically explained by several reasons. One is the 2D nature of the paper-pencil assessment which in fact was more aligned with the studied material in the 2D anatomical atlas group. In a recent study on the effectiveness of a monoscopic 3D visualization technology versus the use of prosected cadaveric specimens, students have performed best on the identification questions aligned with the respective study materials.[47,48] A similar effect has been reported by Henssen and colleagues[49] with the use of cross-sections. Therefore, participants in the 2D anatomical atlas group could have had an advantage over participants in the other two groups. More insight can be gained by future studies that include a combination of assessment methods aligning with each of the interventions.

Another explanation is of a more theoretical nature, namely the unfamiliarity with a new type of 3D visualization technology and the meta-representational competence of students as part of their spatial intelligence. Hegarty has described this competence as the ability to choose the optimal external representation for a task, and effective use of novel external representations, such as interactive visualizations.[50] In their research, novice Navy weather forecasters tended to choose less effective interactive visualization than experts by adding unnecessary visual information to a display in order to interpret a weather forecast.[51] In the current study, relevant 2D images were selected form the anatomical atlas which made it easy for students to identify quickly the useful images. In the intervention group, however, students had to rely on their own choices of visual representations. In an interactive 3D environment, students with lower visual-spatial abilities could therefore be less effective in choosing the right representations of anatomical structures to learn from (e.g., exploring an anatomical structure in the presence of all other structures and/or menu options versus isolating a structure from all other anatomical layers and restricting the user interface to a minimal amount of visual information). Additionally, students with lower VSA tend to use the interactive presentations less effectively. These students for example had difficulties in rotating a digital 3D anatomical structure to a specified view.[50,52] However, with the aid of orientation references, students have been able to successfully manipulate and learn from the virtual model. The tendency to choose a less effective strategy by low performing students has recently been demonstrated by Roach and colleagues in performing a mental rotation task.[53-55] Students with high VSA had a distinct eye-movement pattern in solving mental rotation tasks than low performing students.[53] When low performing students had been instructed by a visual guidance protocol that was based on the eye-movement pattern of high performing students, they had significantly improved in solving the mental rotation tasks.[55] For the reasons stated above, these individual differences can potentially affect the learning strategies of students and are of great interest for further investigation.

Future directions

The findings have implications for both research and education. The modifying effect of VSA should be taken into account when designing new research and analysis strategies, especially in the field of 3D technologies. For educational purposes, stereoscopic 3D AR models have a great potential to be effectively used in small-group teaching settings to stimulate active learning and peer-to-peer interaction by studying a synchronized anatomical 3D models. In addition to traditional ways of teaching, this new teaching tool can be used in the context of personalized learning to meet the students' individual learning needs. Especially, the combination of stereoscopic 3D models and 2D anatomical atlas is worth further research. A possible synergic learning effect would be desirable since the level of anatomical knowledge among medical students still remain insufficient.[1-5] When designing new VR and AR environments one should carefully align the learning environment with the (learning) goals, e.g., VR is better suited for individual learning experiences, whereas AR has many advantages for collaborative and embodied learning.

Limitations

There are some methodological limitations in this study. Firstly, due to the limited availability of hardware, the study was restricted to a maximum of twenty participants in each group. In addition, no distributional data on anatomy knowledge assessment was available beforehand. Therefore, an a priori sample size calculation could not be performed. Only for this reason, a post hoc power analysis was performed based on the observed effect sizes, which turned out to be sufficient. Second concern was the alignment between study materials and assessment. A different form of assessment that is closer to the clinical practice and in line with the learning method (e.g., cadaveric/specimen or digital 3D assessment) should be considered to assess the acquired anatomical knowledge. If not possible, a combination of assessment methods aligning with each of the interventions should be considered. In addition, a long-term retention test would have been valuable to measure the actual retention of anatomical knowledge. Thirdly, the participants were not tested for their lack of depth perception which could be present in about five percent of the study population.[56] Based on these statistics, 1-2 of the twenty participants in the stereoscopic 3D AR group could have perceived the model monoscopically, which could have unfairly lowered the total group score. Lastly, some of the features that were characteristic for the type of intervention, for example hand gestures in stereoscopic 3D AR group and audio cues in both stereoscopic 3D AR and monoscopic 3D desktop groups, could have introduced bias. To eliminate such differences between groups, it is desirable to conduct research within one level of instructional design when possible. Additionally, this will decrease the chance of Hawthorne effect that can occur when learners tend to learn better or harder with a more popular tool or medium, as it could have been the case in the current study.

CONCLUSIONS

Three-dimensional anatomical models that can be viewed stereoscopically in AR can help to optimize anatomical knowledge acquisition and knowledge of students with lower levels of VSA. Further research is needed to identify factors that contribute to the positive learning effect and the most effective way of combining this technology with current education.

REFERENCES

1. McKeown PP, Heylings DJ, Stevenson M, McKelvey KJ, Nixon JR, McCluskey DR. The impact of curricular change on medical students' knowledge of anatomy. Med Educ 2010; 37:954–961.
2. Prince KJ, Scherpbier AJ, Van Mameren H, Drukker J, van der Vleuten CP. Do students have sufficient knowledge of clinical anatomy? Med Educ 2005; 39:326-332.
3. Spielmann PM, Oliver CW. The carpal bones: A basic test of medical students and junior doctors' knowledge of anatomy. Surgeon 2005;3:257–259.
4. Waterston SW, Stewart IJ. Survey of clinicians' attitudes to the anatomical teaching and knowledge of medical students. Clin Anat 2005;18:380–384.
5. Bergman EM, Prince KJ, Drukker J, van der Vleuten CP, Scherpbier AJ. How much anatomy is enough? Anat Sci Educ 2008;1:184–188.
6. Pryde FR, Black SM. Anatomy in Scotland: 20 years of change. Scott Med 2005;50:96–98.
7. Azer SA, Eizenberg N. Do we need dissection in an integrated problem-based learning medical course? Perceptions of first- and second-year students. Surg Radiol Anat 2007;29:173–180.
8. Drake RL, McBride JM, Lachman N, Pawlina W. Medical education in the anatomical sciences: The winds of change continue to blow. Anat Sci Educ 2009;2:253–259.
8. Klatzky RL, Lederman SJ. Haptic object perception: Spatial dimensionality and relation to vision. Philos Trans R Soc Lond B Biol Sci 2011;366:3097–3105.
9. Bergman EM, de Bruin AB, Herrler A, Verhrijen IW, Scherpbier, van der Vleuten CP. Students' perceptions of anatomy across the undergraduate problem-based learning medical curriculum: A phenomenographical study. BMC Med Educ 2013;13:152–162.
9. Reid S, Shapiro L, Louw G. How haptics and drawing enhance the learning of anatomy. Anat Sci Educ 2018;12:164–172.
10. Yammine K, Violato C. A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy. Anat Sci Educ 2015;8:525–538.
11. Garg A, Norman GR, Spero L, Maheshwari P. Do virtual computer models hinder anatomy learning? Acad Med 1999;74:S87–S89.
12. Garg A, Norman G, Spero L, Taylor I. Learning anatomy: Do new computer models improve spatial understanding? Med Teach 1999;21:519–522.
13. Garg AX, Norman GR, Eva KW, Spero L, Sharan S. Is there any real virtue of virtual reality? The minor role of multiple orientations in learning anatomy from computers. Acad Med 2002;77:S97–S99.
14. Levinson AJ, Weaver B, Garside S, McGinn H, Norman GR. Virtual reality and brain anatomy: a randomised trial of e-learning instructional designs. Med Educ 2007;41:495–501.
15. Naaz F. Learning from graphically integrated 2D and 3D representations improves retention of neuroanatomy. University of Louisville: Louisville, KY. Doctorate of Philosophy Dissertation. 2002. 76 p.
16. Huk T. Who benefits from learning with 3D models? The case of spatial ability. J Comput Assist Learn 2006;22:392–404.
17. Khot Z, Quinlan K, Norman GR, Wainman B. The relative effectiveness of computer-based and traditional resources for education in anatomy. Anat Sci Educ 2013;6:211–215.
18. Cui D, Wilson TD, Rockhold RW, Lehman MN, Lynch JC. Evaluation of the effectiveness of 3D vascular stereoscopic models in anatomy instruction for first year medical students. Anat Sci Educ 2017;10:34–45.

19. Luursema JM, Verwey WB, Kommers PA, Geelkerken RH, Vos HJ. Optimizing conditions for computer-assisted anatomical learning. Interact Comput 2006;18:1123–1138.
20. Luursema JM, Verwey WB, Kommers PA, Annema JH. The role of stereopsis in virtual anatomical learning. Interact Comput 2008;20:455–460.
21. Luursema JM, Vorstenbosch M, Kooloos J. Stereopsis, visuospatial ability, and virtual reality in anatomy learning. Anat Res Int 2017:1493135.
22. Hackett M, Proctor M. The effect of autostereoscopic holograms on anatomical knowledge: A randomized trial. Med Educ 2018;52:1147–1155.
23. Wainman B, Wolak L, Pukas G, Zheng E, Norman GR. The superiority of three-dimensional physical models to two-dimensional computer presentations in anatomy learning. Med Educ 2018; 52:1138–1146.
24. Jolicoeur P, Milliken B. Identification of disoriented objects: effects of context of prior presentation. J Exp Psychol Learn Mem Cogn 1989;15:200–210.
25. Kourtzi Z, Erb M, Grodd W, Bulthoff HH. Representation of the perceived 3-D object shape in the human lateral occipital complex. Cereb Cortex 2003;13:911–920.
26. Verhoef B-E, Vogels R, Janssen P. Binocular depth processing in the ventral visual pathway. Philos Trans R Soc Lond B Biol Sci 2016;371:20150259.
27. Cook DA. The research we still are not doing: An agenda for the study of computer-based learning. Acad Med 2005;80:541–548.
27. Moro C, Štromberga Z, Raikos A, Stirling A. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. Anat Sci Educ 2017;10:549–559.
28. Kuehn BM. Virtual and augmented reality put a twist on medical education. JAMA 2018;319:756–758.
29. Küçük S, Kapakin S, Göktaş Y. Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. Anat Sci Educ 2016; 9:411–421.
30. Barmaki R, Yu K, Pearlman R, Shingles R, Bork F, Osgood GM, Navab N. Enhancement of anatomical education using augmented reality: An empirical study of body painting. Anat Sci Educ 2019;12:599–609.
31. Sugiura A, Kitama T, Toyoura M, Mao X. The use of augmented reality technology in medical specimen museum tours. Anat Sci Educ 2019;12:561–571.
32. Cook DA, Beckman TJ. Reflections on experimental research in medical education. Adv Health Sci Educ Theory Pract 2010;15:455–464.
32. McJunkin JL, Jiramongkolchai P, Chung W, Southworth M, Durakovic N, Buchman CA, Silva JR. Development of a mixed reality platform for lateral skull base anatomy. Otol Neurotol 2018;39:e1137–e1142.
33. Putz R, Pabst R. Sobotta Atlas of Human Anatomy. Part 2. 3rd Ed. Houten, The Netherlands: Bohn Stafleu van Loghum. 2006; 399 p.
34. Moore KL, Dalley AF, Agur AM. Clinically Oriented Anatomy. 7th Ed. Philadelphia, PA: Lippincott Williams & Wilkins. 2013;1168 p.
35. Biggs J. Enhancing teaching through constructive alignment. High Educ 1996;32:347–364.
36. Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR. Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain. 1st Ed. New York, NY: David McKay Company. 1956. 207 p.
37. Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. Percept Mot Skills 1978;47:599–604.

38. Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C. A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. Brain Cognit 1995;28:39–58.
39. Guillot A, Champely S, Batier C, Thiriet P, Collet C. Relationship between spatial abilities, mental rotation and functional anatomy learning. Adv Health Sci Educ Theory Pract 2007;12:491–507.
40. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities and anatomy knowledge assessment: A systematic review. Anat Sci Educ 2017;10:235–241.
41. Ekstrom RB, French J, Harman HH, Dermen D.Kit of Factor-Referenced Cognitive Tests. Princeton, NJ: Educational Testing Service. 1976. 314 p.
42 Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Ed. Hillsdale, NJ: Lawrence Earlbaum Associates. 1988. 400 p.
43. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. 3rd Ed. Philadelphia, PA: Lippincott Williams and Wilkins. 2008. 758 p.
44. Corraini P, Olsen M, Pedersen L, Dekkers OM, Vandenbroucke JP. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. Clin Epidemiol 2017;9:331–338.
45. Bulthoff HH, Edelman SY, Tarr MJ. How are three-dimensional objects represented in the brain? Cereb Cortex 1995;5:247–260.
46. Mayer RE. Cognitive theory of multimedia learning. In: Mayer RE (Editor). Multimedia Learning. 2nd Ed. Santa Barbara, CA: Cambridge University Press. 2014. p 43–47.
47. Mitrousias V, Karachalios TS, Varitimidis SE, Natsis K, Arvanitis DL, Zibis AH. Anatomy learning from prosected cadaveric specimens versus plastic models: A comparative study of upper limb anatomy. Anat Sci Educ 2020;13:436-444.
48. Mitrousias V, Varitimidis SE, Hantes ME, Malizos KN, Arvanitis DL, Zibis AH. 2018. Anatomy learning from prosected cadaveric specimens versus three-dimensional software: A comparative study of upper limb anatomy. Ann Anat 218:156–1564.
49. Henssen DJ, van den Heuvel L, De Jong G, Vorstenbosch MA, van Cappellen van Walsum AM, Van den Hurk MM, Kooloos JG, Bartels RH. 2020. Neuroanatomy learning: Augmented reality vs. cross-sections. Anat Sci Educ (in pres; doi 10.1002/ase.1912).
50. Hegarty M. Chapter 7 - Components of spatial intelligence. Psychol Learn Motiv 2010;**52**:265–297.
51. Smallman HS, Hegarty M. Expertise, spatial ability and intuition in the use of complex visual displays. In: Proceedings of the 51st Annual Meeting of the Human Factors and Ergonomics Society (HFES 2007); Baltimore, MD, 2007 October 1-5. p 2000–2004. Human Factors and Ergonomics Society: Santa Monica, CA.
52. Stull AT, Hegarty M, Mayer RE. Orientation references: Getting a handle on spatial learning. J Educ Psychol 2009;101:803–816.
53. Roach VA, Fraser GM, Kryklywy JH, Mitchell DG V, Wilson TD. Different perspectives: Spatial ability influences where individuals look on a timed spatial test. Anat Sci Educ 2017;10:224–234.
54. Roach VA, Fraser GM, Kryklywy JH, Mitchell DG, Wilson TD. Time limits in testing: An analysis of eye movements and visual attention in spatial problem solving. Anat Sci Educ 2017;10:528–537.
55. Roach VA, Fraser GM, Kryklywy JH, Mitchell DG, Wilson TD. Guiding low spatial ability individuals through visual cueing: The dual importance of where and when to look. Anat Sci Educ 2019;12:32–42.
56. Mather G. Foundations of Perception. 1st Ed. Hove, UK: Psychology Press. 2006. 400 p.