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

Bogomolova, K.

4

The effect of binocular disparity on learning anatomy with stereoscopic augmented reality visualization: a double-center randomized controlled trial

Katerina Bogomolova, Marc A.T.M. Vorstenbosch, Inssaf El Messaoudi, Micha Holla, Steven E. R. Hovius, Jos A. van der Hage, Beerend P. Hierck

Accepted for publication in Anatomical Sciences Education

ABSTRACT

Background

Binocular disparity provides one of the important depth cues within stereoscopic three-dimensional (3D) visualization technology. However, there is limited research on its exact effect on learning within 3D augmented reality (AR) environment. This study evaluated the effect of binocular disparity on acquisition of anatomical knowledge and perceived cognitive load in relation to visual-spatial abilities (VSA).

Methods

In a double-center randomized controlled trial, first-year (bio)medical undergraduates studied lower extremity anatomy in an interactive 3D AR environment either with a stereoscopic 3D view (n = 32) or monoscopic 3D view (n = 34). VSA were tested with a Mental Rotation Test (MRT). Anatomical knowledge was assessed by a validated 30-item written test and 30-item specimen test. Cognitive load was measured by the NASA-TLX questionnaire.

Results

Students in the stereoscopic 3D and monoscopic 3D groups performed equally well in terms of percentage correct answers (written test: 47.9 ± 15.8 vs 49.1 ± 18.3; $p = .635$; specimen test: 43.0 ± 17.9 vs 46.3 ± 15.1; $p = .429$), and perceived cognitive load scores (6.2 ± 1.0 vs 6.2 ± 1.3; $p = .992$). Regardless of intervention, VSA were positively associated with the specimen test scores ($\eta^2 = 0.13$, $p = .003$), perceived representativeness of the anatomy test questions ($p = .010$) and subjective improvement in anatomy knowledge ($p < .001$).

Conclusions

In conclusion, binocular disparity does not improve learning anatomy. Motion parallax should be considered as another important depth cue that contributes to depth perception during learning in a stereoscopic 3D AR environment.

INTRODUCTION

Three-dimensional visualization technology (3DVT) has a great potential to contribute to a better learning and understanding of anatomy in medical education. [1,2] Its contribution is becoming necessary in times of decreased teaching hours of anatomy and exposure to traditional teaching methods, such as cadaveric dissections.[3-5] Additionally, anatomical knowledge is reported to be insufficient among medical students and junior doctors, who still experience difficulties in translating the acquired knowledge into clinical practice.[6-9] However, to know *that* 3DVT can be highly effective, is currently not enough. There is a need to know *how* this technology works to be able to implement it in everyone's unique educational setting.[10]

Stereoscopic versus monoscopic three-dimensional visualization technology

In real life, stereoscopic vision is obtained due to positioning of the human eyes in a way that generates two slightly different retinal images of an object, also referred to as binocular disparity.[11] The same effect can be mimicked within 3DVT by presenting a slightly shifted and rotated image to the right and left eye. Stereoscopic vision can be obtained by supportive devices such autostereoscopic displays e.g., Alioscopy 3D Display (Alioscopy, Paris, France), anaglyphic or polarized glasses, or by a head-mounted display e.g., HoloLens™ (Microsoft Corp., Redmond, WA), or Oculus Rift™ (Oculus VR, Menlo Park, CA) and HTC VIVE™ (High Tech Computer Corp., New Taipei City, Taiwan). Hololens™ is used to create interactive augmented reality (AR), also referred to as mixed reality (MR). Oculus Rift™ and HTC VIVE™ are predominantly used to create virtual reality (VR) environments. A binocular vision of the viewer, though, is required to perceive the obtained visual depth. In the absence of stereoscopic vision, 3D effect is mimicked by monocular cues, such as shading, coloring, relative size and motion parallax.[12] The examples of monoscopic 3DVT include 3D anatomical models that can be explored from different angles on a computer, tablet or phone.[13]

Distinction between stereoscopic and monoscopic modalities within 3DVT is essential to make since different processes are involved. Research has shown that recognition of digital 3D objects appears to be greater when objects are presented stereoscopically.[14-17] More importantly, the type of modality can significantly affect learning. Monoscopic 3DVT has been demonstrated to have disadvantages for students with lower VSA.[18-23] The disadvantages are explained by the ability-as-enhance hypothesis within the cognitive load theory.[24,25] Initially, it has been hypothesized that 3D objects are remembered as key view-based 2D images.[20,21,26]

Consequently, when an unfamiliar 3D object is viewed from multiple angles, an increase in cognitive load occurs while generating a proper mental representation of a 3D object. During this process, individuals with higher visual-spatial abilities are able to devote more cognitive resources to building mental connections, while students with low VSA get cognitively overloaded.[20,26] The latter leads to underperformance among students with lower VSA. However, as research has shown, with stereoscopic 3DVT, students with lower VSA are able to achieve comparable levels of performance of students with higher VSA.[28] This can be explained by the fact that the mental 3D representations of the object are already built and provided by the stereoscopic projection and perception. Consequently, mental steps, that are required to build a mental 3D representation, can be skipped while leaving a sufficient amount of cognitive resources. In this way, students with lower VSA are able to allocate these resources to learning.

The role of stereopsis

In health care, the benefits of stereoscopic visualization within 3D technologies have been recognized for years.[29-33] Development and utilization of stereoscopic 3DVT is still growing, especially in the surgical field. Several examples include preoperative planning and identification of tumor with stereoscopic AR.[31,34,35] Another examples include stereoscopic visualization during minimal invasive surgeries where stereoscopic view of the surgical field would improve spatial understanding and orientation during laparoscopic surgeries.[29,36] Stereoscopic visualization even showed to shorten operative time of laparoscopic gastrectomy by reducing the intracorporeal dissection time.[37]

The beneficial effect of stereopsis on learning anatomy has been recently demonstrated in a comprehensive systematic review and meta-analysis.[38] In the meta-analysis, the comparisons between studies were made within a single level of instructional design, e.g., stereopsis was isolated as the only true manipulated element in the experimental design. The positive effect of stereopsis was demonstrated across different types of 3D technologies combined, predominantly using the VR headsets and 3D shutter glasses for desktop applications. How learning experience is affected by a particular type of stereoscopic 3DVT, remains a topic for further exploration.

Stereoscopic augmented reality in anatomy education

Stereoscopic AR is a new generation of 3DVT technology that combines stereoscopic visualization of 3D computer-generated objects with the physical environment. The main distinguishing feature from other types of AR is the ability to provide stereoscopic vision, e.g., to perceive the anatomical model in real 3D. Additionally, it provides the ability to walk around the model and explore it form all possible angles without losing the sense of the user's own environment. This view can be obtained with e.g., HoloLens®, a head-mounted

display from Microsoft (*Supplementary material 2*). In the previous study, authors evaluated the effectiveness of stereoscopic 3D AR visualization in learning anatomy of the lower leg among medical undergraduates.[23] Learning with a stereoscopic 3D AR model was more effective than learning with a monoscopic 3D desktop model. Interestingly, the observed positive learning effect was only present among students with lower VSA. Stereoscopic vision was hypothesized to be one of the distinguishing features of intervention modality that could explain these differences. However, since the comparisons were made within different levels of instructional design, e.g., stereoscopic vision was not isolated as the only true manipulated element, the actual effect of stereoscopic vision remained unrevealed. A similar study design approach was used by Moro and colleagues who compared the effectiveness of HoloLens with mobile-based AR environment.[39] Although both learning modes were effective in terms of acquired anatomical knowledge, comparisons were still made within different levels of instructional design.

In another recent study of the role of stereopsis in 3DVT, Wainman and colleagues have isolated binocular disparity by covering the non-dominant eye of participants. Authors reported positive effect of stereoscopic vision in VR, but not in AR.[40] Although it was a simple and vivid way of isolating stereopsis, participants in the control group remained aware of their condition which could have influenced the outcomes. Additionally, different effect measures of stereopsis in VR and AR suggest that the type of technology is decisive for the learning effect caused by stereoscopic vision.

Objectives and aims

Based on considerations described above and lessons learned from previous research, this study aimed to evaluate the role of binocular disparity in a stereoscopic 3D AR environment within a single level of instructional design. Therefore, the primary objective was to evaluate whether learning with a stereoscopic view of a 3D anatomical model of the lower extremity was more effective than learning with a monoscopic 3D view of the same model among medical undergraduates. The secondary objectives were to compare the perceived cognitive load among groups, and to evaluate whether VSA would modify the outcomes.

Authors hypothesized that learning within a stereoscopic 3D AR environment would be more effective than learning within a monoscopic 3D AR environment. Authors also hypothesized that the perceived cognitive load in the stereoscopic 3D view group would be lower than in the monoscopic 3D view group, and that the students with lower VSA would benefit most from the stereoscopic 3D view of the model.

MATERIALS AND METHODS

Study design

A single-blinded double-center randomized controlled trial was conducted at the Leiden University Medical Center (LUMC) and the Radboudumc University Medical Center (Radboudumc), the Netherlands. The study was conducted within a single level of instructional design, e.g., isolating binocular disparity as the only true manipulated element. (Figure 1). The study was approved by the Netherlands Association for Medical Education (NVMO) Ethical Review Board (NERB case number: 2019.5.8).

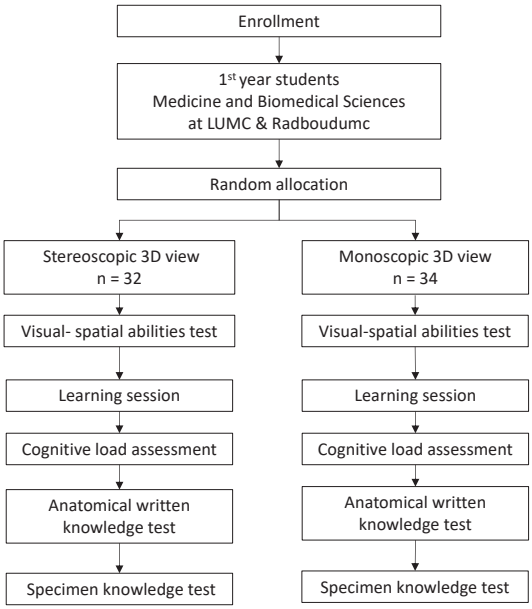

Figure 1. Flowchart of study design. LUMC, Leiden University Medical Center; Radboudumc, Radboud University Medical Center; 3D, three-dimensional; n, number of participants.

Study population

Participants were first-year undergraduate students of Medicine and Biomedical Sciences with no prior knowledge of the lower extremity anatomy. The baseline knowledge was not assessed to avoid extra burden for students and possible influence on learning during the intervention and performance on the post-tests. [41] Participation was voluntary and written consent was obtained from all participants. Participation did not interfere with the curriculum and the assessment results did not affect student's academic grades. Participants received a financial compensation at the completion of the experiment.

Randomization and blinding of participants

Participants were randomly allocated to either stereoscopic 3D view or monoscopic 3D view groups using an Excel Random Group Generator (Microsoft Excel for Office 365 MSO, version 2012). Participants were not aware of the distinction between stereoscopic and monoscopic 3D views and remained blinded to the type of condition during the entire experiment. The intended goal of the study and individual allocation to study arms was clarified and debriefed directly after experiment.

Educational interventions

An interactive AR application *DynamicAnatomy* for Microsoft HoloLens®, version 1 (Microsoft Corp., Redmond, WA) was developed at the department of Anatomy at Leiden University Medical Center and the Centre for Innovation of Leiden University. The application represented a dynamic and fully interactive stereoscopic 3D model of the lower extremity. Users perceived the 3D model as a virtual object in their physical space without losing the sense of their own physical environment. The object centered view, i.e., dynamic exploration, enabled learners to walk around the model and explore it from all possible angles. Active interaction included size adjustments, showing or hiding anatomical structures by group or individually, visual and auditory feedback on structures and anatomical layers, and animation of the ankle movements. The anatomical layers included musculoskeletal, connective tissue, and neuro-vascular systems. During this experiment, study participants studied the musculoskeletal system. Prior to the experiment, participants completed a 10-minutes training module (without anatomical content) to get familiar with the use of application and device.

In the intervention group, the 3D model of the lower extremity was presented and perceived stereoscopically as intended by the supportive AR device. In the control group, binocular disparity was eliminated technically by projecting an identical, i.e., non-shifted and non-rotated, image to both eyes. This adjustment resulted in a monoscopic view of the identical 3D anatomical model. Therefore, binocular disparity was isolated as the only true manipulated element in this experimental design. All other features of the AR application described above remained available and identical in both conditions.

Stereovision of participants

Stereovision of participants was measured by a Random Dot 3 - LEA Symbols® Stereoacuity Test (Vision Assessment Corp., Elk Grove Village, IL) prior to the experiment to identify individuals with absent stereovision.

Baseline characteristics

Informed consent and baseline questionnaire were administered prior to the start of the experiment.

Assessment of VSA

VSA were assessed prior to the start of the learning session. Mental visualization and rotation, as the main components of visual-spatial abilities, were assessed by the 24-item Mental Rotation Test (MRT), previously validated by Vandenberg and Kuse and redrawn by Peters and colleagues.[42,43] This psychometric test is being widely used in the assessment of VSA and has repeatedly shown its positive association with anatomy learning and assessment.[44,45] The post-hoc level of internal consistency (Cronbach's alpha) of the MR test in this study was 0.94. Duration of this test was ten minutes without intervals.

Learning session

Participants received a handout with a description of the learning goals and instructional activities. The development of learning goals and instructions was based on the constructive alignment theory to ensure alignment between the intended learning outcomes, instructional activities and knowledge assessment (*Supplementary material 4A, 4B*).[23] Learning goals were formulated and organized according to Bloom's Taxonomy of Learning Objectives.[46] An independent expert 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, understanding the function of muscles based on their origin and insertion, and location and organization of these structures in relation to each other. Duration of the learning session was 45 minutes.

Cognitive load assessment

Cognitive load was measured by the validated NASA-TLX questionnaire immediately after the session.[47] The NASA-TLX questionnaire is a subjective, multidimensional assessment instrument for perceived workload of task, in this case the workload required to study the anatomy of lower extremity. The items included mental demand, physical demand, temporal demand, performance, effort, and frustration level. Response options ranged from low (0 point) to high (10 points). The total score was calculated according to the prescriptions of the questionnaire and ranged also between 0 and 10 points.

Written anatomy knowledge test

A previously validated 30-item knowledge test consisted of a combination of 20 extended matching questions and 10 open-ended questions (*Supplementary material 6*).[23] Anatomical 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. Content validation was assessed by two experts in the field of anatomy and plastic and reconstructive surgery. The test was piloted among 12 medical students for item clarity. The level of internal consistency (Cronbach's alpha) was 0.78. Duration of assessment was 30 minutes.

Specimen knowledge test

Plastinated specimen test covered a total of 30 anatomical structures on 12 specimens distributed over 10 stations (*Supplementary material 7)*. Content validation was assessed by one expert in the field of anatomy. Each station included 3-4 structures that were labeled on one or more specimen. Participants were asked to name the labeled structures or to indicate what movement is initiated by a particular structure. The post-hoc level of internal consistency (Cronbach's alpha) of the test was 0.90. Duration of this assessment was 20 minutes with a maximum of two minutes per station.

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 an independent t-test for differences in means and X^2 test for differences in proportions. Anatomical knowledge was defined as mean percentage of correct answers on the written knowledge test and the specimen test. Cognitive load was defined as the mean score on the NASA-LTX questionnaire. Differences in outcome measures between groups were assessed with an independent t-test. Additionally, a ANCOVA was performed to measure the effect of the intervention for different levels of visual-spatial abilities by including the interaction term "MRT score x intervention" in the model. MRT score was also included as a covariate to measure its effect on outcomes regardless of intervention. Additional analyses were performed for sex differences. 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 66 students were included (Table 1). All participants were able to perceive spatial visual depth as measured by the stereoacuity test.

Table 1. Baseline characteristics of included participants.

	Stereoscopic 3D view	Monoscopic 3D view	p value
	n = 32	n = 34	
Sex, n (%)			
Male	11 (34.4)	16 (47.1)	.295
Female	21 (65.6)	18 (52.9)	
Age, mean ±SD, y	19.2 ± 1.3	19.0 ± 1.9	.754
Medical center, n (%)			
Leiden University MC	16 (50.0)	16 (47.1)	.811
Radboudumc University MC	16 (50.0)	18 (52.9)	
Study, n (%)			
Medicine	30 (93.8)	33 (97.1)	.519
Biomedical sciences	2 (6.3)	1 (2.9)	
Videogame, n (%)			
Never	23 (71.9)	22 (64.7)	.397
0-2 hours a week	6 (18.8)	6 (17.6)	
2-10 hours a week	2 (6.3)	6 (17.6)	
>10 hours a week	1 (3.1)	0 (0.0)	
AR experience before, n (%)			
No	26 (81.3)	23 (67.6)	.207
Yes	6 (18.8)	11 (32.4)	
Mental Rotation Test, mean ±SD	14.1± 5.1	15.7 ± 5.3	.212

Minimal and maximal scores range between 0-24 for the Mental Rotation Test, 3D, three-dimensional; AR, Augmented Reality; n, number of participants; SD, standard deviation; y, years; MC, medical center.

As shown in Figure 2, participants in the stereoscopic 3D view group performed equally well as the participants in the monoscopic 3D view group on the written knowledge test (47.9 ± 15.8 vs 49.1 ± 18.3; p = .635). Likewise, no differences were found for each knowledge domain separately (*factual:* 34.1 ± 19.5 vs 34.3 ± 19.0; p = .970; *functional:* 33.4 ± 16.4 vs 31.5 ± 13.7; p = .611; *spatial:* 50.4 ± 15.2 vs 47.3 ± 13.5; p = .384).

Percentages correct answers on the specimen test were not significantly different between groups (43.0 ± 17.9 vs 46.3 ± 15.1; p = .429) (Figure 2).

The observed similarities between groups on the knowledge tests were reflected by the cognitive load scores that were similar in both groups (6.2 ± 1.0 vs 6.2 ± 1.3; p = .992) (Figure 2).

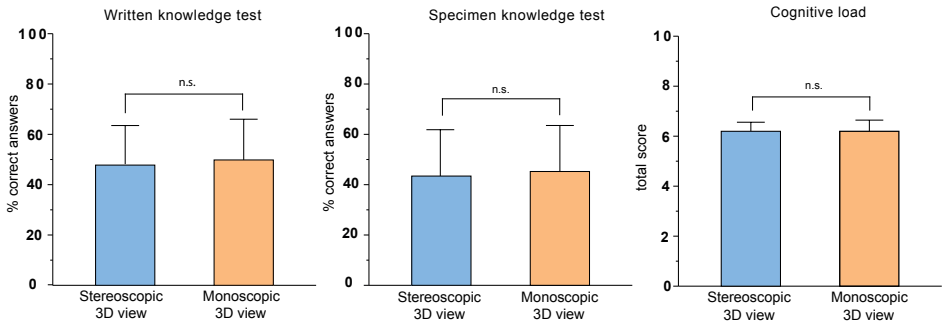

Figure 2. Differences in overall mean percentages correct answers on the written knowledge test, specimen knowledge test and cognitive load test between stereoscopic 3D view (n = 32) and monoscopic 3D view (n = 34) groups. 3D, three-dimensional; n.s., not significant.

As shown in Table 2, there were no significant differences in learning experience between stereoscopic 3D view and monoscopic 3D view groups. All participants enjoyed studying (4.4 ± 0.7 vs 4.3 ± 0.8; p = .492) and reported an improved anatomical knowledge of the lower extremity (4.2 ± 0.9 vs 4.1 ± 0.7; p = .502). Five versus four participants reported the device to be heavy on their nose after a longer period of study time in stereoscopic and monoscopic 3D groups respectively (p = .794). Headache and nausea were reported by one participant in the stereoscopic 3D group.

Table 2. Differences in learning experience between groups.

	Stereoscopic 3D view	Monoscopic 3D view	p value
	n = 32	n = 34	
The study time was long enough to study the required number of anatomical structures	2.4 ± 1.0	2.5 ± 1.1	.500
The questions in anatomy test were representative for the studied material	3.7 ± 1.0	3.4 ± 1.1	.249
My knowledge about anatomy of the lower leg is improved after studying	4.2 ± 0.9	4.1 ± 0.7	.502
Learning material was easy to use	3.5 ± 0.9	3.4 ± 0.9	.378
I enjoyed studying	4.4 ± 0.7	4.3 ± 0.8	.492
I would recommend studying with … to my fellow students	4.1 ± 0.8	3.7 ± 0.9	.057

Response options on a five-point Likert scale ranged from 1 = very dissatisfied to 5 = very satisfied. Scores are expressed in means ± SD. SD, standard deviation; 3D, three-dimensional; n, number of participants.

The effect of VSA

In both study groups, mean scores on the written knowledge test and for each knowledge domain separately remained similar for all levels of MRT scores, as measured by the interaction term in the ANCOVA analysis (written knowledge test: $F(1,62) = 0.51$, $p = .393$; *factual:* $F(3,62) = 0.15$, $p = .925$; *functional:* $F(3,62) = 1.04$, $p = .381$; *spatial:* $F(3,62) = 0.92$, $p = .435$).

Similar effects were found for the specimen knowledge test ($F(1,62) = 0.00$, $p = .998$). However, regardless of intervention, MRT scores were significantly and positively associated with the specimen test scores, as shown in Figure 3 ($F(1,62) = 9.37$, Partial $\eta 2 = 0.13$, $p = .003$).

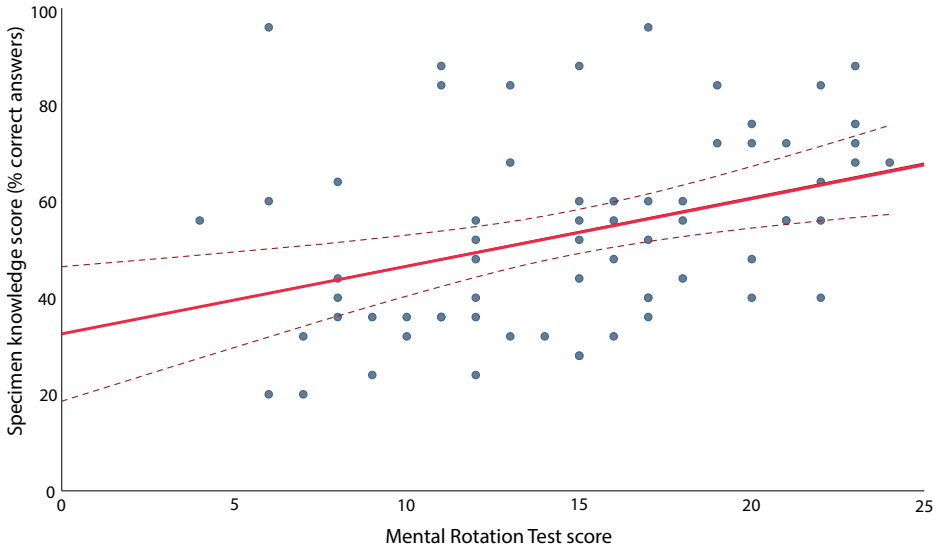

Figure 3. Relationship between Mental Rotation Test (MRT) scores and specimen test scores. A regression analysis graph illustrating a positive association between visual-spatial abilities and specimen test scores. MRT, Mental Rotations Test.

The perceived cognitive load scores remained similar for all levels of VSA in both study groups ($F(1,62) = 2.26$, $p = .138$). Regardless of intervention, MRT scores were not associated with the perceived cognitive load scores.

ANCOVA analysis for learning experience revealed that participants in the monoscopic 3D view group found the anatomy test questions significantly less representative for the studied material than participants in the stereoscopic 3D view group. This difference was only present among individuals with lower visual-spatial abilities scores ($F(1,62) = 2.26$, $p = .044$). As independent variable, VSA scores were significantly and positively associated with the perceived representativeness of the anatomy test questions ($p = .010$) and subjective improvement in anatomy knowledge of the lower extremity ($p < .001$).

Sex differences

On baseline, males achieved significantly higher MRT scores than females (17.5 ± 4.9 vs 13.2 vs 4.9; $p = .001$). Both sexes performed equally well on written anatomical knowledge test (written knowledge test: 52.4 ± 18.9, $p = .96$; *factual:* 37.9 ± 20.4 vs 31.5 ± 17.9, $p = .180$; *functional:* 36.3 ± 17.4 vs 29.9 ± 12.7, $p = .091$; *spatial:* 51.4 ± 14.6 vs 47.1 ± 14.1, $p = .242$). However, males achieved significantly higher scores on the specimen test (51.5 ± 15.8 vs 40.0 ± 15.6; $p = .005$). Perceived cognitive load remained similar for both sexes (6.2 ± 1.2 vs 6.1 ± 1.1, $p = .915$).

DISCUSSION

This study evaluated the effect of binocular disparity on learning anatomy in a stereoscopic 3D AR environment. Against author's expectations, no differences were found between stereoscopic 3D and monoscopic 3D view groups in terms of acquired anatomical knowledge and perceived cognitive load during learning. VSA, however, were significantly and positively associated with practical anatomical knowledge regardless of intervention. Additionally, VSA were positively associated with the perceived representativeness of anatomy test questions and the subjective improvement in anatomy knowledge of the lower extremity.

Although binocular disparity is generally considered to provide one of the important depth cues in 3D visualization, its exclusive effect on learning and cognitive load was revealed to be not significant in a stereoscopic 3D AR environment. To the author's knowledge only one study, performed by Wainman and colleagues, has evaluated the role of binocular disparity within the same type of technology.[40] Likewise, Wainman and colleagues found no beneficial effect of stereopsis on learning. The only difference between both studies was the way binocular disparity was eliminated. While in the current study a monoscopic view was obtained technically by presenting identical images to both eyes, Wainman and colleagues achieved monocular view by closing the dominant eye of participants with a patch. In addition, Wainman and colleagues have compared the effect of binocular disparity in AR to its effect in VR. The effect of stereoscopic vision in VR appeared to be significantly greater than in AR. In fact, learning with a stereoscopic 3D model in AR was less effective than in VR. This effect was explained by various degrees of stereopsis that different types of technologies can generate.

On the other hand, the findings suggest that other important depth cues could have compensated for the absence of stereopsis. During the experiment participants were able to walk around the 3D anatomical model and explore the model from all possible angles which is unique for a stereoscopic AR environment. This type of dynamic exploration, also referred to as motion parallax, is able to provide strong depth information.[48] Additional literature searches in the field of neurosciences education revealed that motion parallax in some cases can be even more effective than binocular disparity alone.[49-51] More interestingly, an interaction between both depth cues can exist.[52] For instance, subjects were asked to perform series of explorative tasks under three depth cue conditions: binocular disparity, motion parallax and combination of both depth cues.[50] The combination of binocular disparity and motion parallax resulted in an equal amount of correct answers as did the motion parallax condition (84% vs 80%; $p = .231$). However, in the absence of motion parallax, binocular disparity condition

contributed to significantly less correct answers (60% vs. 80%; $p < .001$). In another study, that motion parallax improved performance in recovering 3D shape of objects in a monoscopic view, but not in a stereoscopic view.[53] Therefore, motion parallax could have reasonably compensated for the absence of binocular disparity and generated a sufficient 3D perception of the monoscopically projected 3D model. Further research is needed to evaluate to what extent motion parallax, alone and in combination with binocular disparity, affects learning.

Another effect of dynamic exploration, that could have occurred during this experiment, is the embodied cognition on learning.[54-56] Previous research has shown that using gestures and body movements helps students acquire anatomical knowledge. For instance, students who have engaged in miming using representational and metaphorical gestures while learning functions of central nervous system, have improved their marks with 42% in comparison with didactic learning.[56] Similar concept applies for mimicking specific joint movements in order to memorize them and being able to recall the structures names and to localize them on a visual representation.[55] Students in the current experiment were also using gestures while dissecting the anatomical layers and structures. That could have helped them memorizing structures while using similar gestures again and again. Additionally, students tended to move their own leg in a synchronized manner with the animated 3D model. Such engagement could have resulted in embodied learning and contribute to better learning within both modalities.

The effect of VSA

In the current study anatomical knowledge was tested both by written and practical examinations. Both assessment methods were chosen to ensure a better alignment between learning and assessment of spatial knowledge. Consistent with previous research, VSA were positively associated with anatomical knowledge as measured by the practical specimen test.[45,57] However, VSA did not modify the observed outcomes as expected. In other words, individuals with lower VSA did not show different trajectory of learning with either monoscopic or stereoscopic 3D views of the model. Also, they did not experience significant differences in perceived cognitive load. This is in contrast with previous body of evidence on an aptitude-treatment effect caused by VSA when learning with different types of 3DVT.[23,28,58] If motion parallax was reasonably compensating for the absence of binocular disparity, as discussed above, then it does explain why students with lower VSA performed equally well within both conditions. These individuals were still able to generate proper 3D mental representations of the model within the monoscopic 3D view group and experienced equal amount of cognitive load during learning. Although the modifying effect of VSA on objective outcomes was not observed in current study, it

was affecting the subjective outcomes regarding learning experience. This is particularly interesting, since the monoscopic 3D group with low VSA found the practical assessment items to be less representative of their learning environment than the stereoscopic 3D group.

Another explanation for the absence of modifying effect of VSA lies within the scale of spatial abilities needed for the task at hand. For spatial abilities a division between small- and large-scale space can be made, with small scale referring to space within arm's length, e.g., tabletop tasks. Large scale space refers to when locomotion is needed to interact with the spatial environment. As participants were walking around the model, large scale spatial processing takes place. As previously shown, a partial dissociation is found for small- and large-scale spatial abilities.[59] It could therefore be that the small-scale task of mental rotation used here, may not substantially relate to the large-scale spatial task of interacting with the model. Alternatively, large scale spatial tests, especially those relying on perspective taking, could show the interaction with task performance as hypothesized here.

Lastly, the observed sex differences in visual-spatial abilities scores in favor of males are in line with previous research.[60-64] More interestingly, males significantly outperform females on the specimen test, but not on the written knowledge test. Again, these findings suggest that the practical examination questions rely on visual-spatial abilities skills more than written knowledge test questions do. This is further supported by the work of Langlois and colleagues who have reviewed relationship between visual-spatial abilities test and anatomy knowledge assessment.[45] Authors have found significant relationship between spatial abilities test and anatomy knowledge assessment using practical examination, while relationship between spatial abilities an spatial multiple-choice questions remained unclear. Therefore, both findings suggest that practical examination questions are more reliable in testing spatial anatomical knowledge than multiple-choice questions, even when designed properly. Further research is needed to explore how spatial multiple-choice questions are mentally processed during examination in comparison to practical examination questions.

Limitations

This study has several strengths and limitations. To authors' knowledge, this was the first single blinded randomized controlled trial to evaluate the effect of binocular disparity on learning in 3D AR environment within two academic centers and within one single level of instructional design. Along with the validated measurement instruments, it has maximized the internal and external validity of the results. On the other hand, participation was voluntary, and a selection bias could occur. The results could have been different if measured within the entire students' population. However, the baseline VSA scores among current study sample bear strong resemblance to VSA scores of the entire cohort of first-year medical undergraduates (14.9 vs. 14.4), as measured previously by Vorstenbosch and colleagues.[65] Another limitation was the relatively small sample size. Due to the limited availability of devices, authors were restricted to a maximum number of participants. It is possible that a much larger sample size could have revealed significant differences between interventions. The possible compensating effect of motion parallax and the effect of large-scale spatial abilities can also be considered as potential confounders that have influenced the internal validity. These new insights can help reveal the exact effect of both factors on learning. It is also important to note that the authors choose to not assess baseline knowledge to avoid extra burden for students and possible influence on learning during the intervention and performance on the post-tests. In this way any differences in prior knowledge that could have been present among students were not taken into account. Lastly, spatial knowledge questions in this study were carefully designed to stimulate mental visualizations skills. However, these questions can still be processed without spatial reasoning or just being best guessed when questions get too difficult to answer. Consequently, stereoscopic visualization of anatomy would not be that helpful in processing these type of questions.

Future implications

The findings of this study have implications for both research and education. As stated previously, the aptitude treatment interaction caused by VSA should be taken into account when designing new research, especially when evaluating 3D technologies and their effect on learning. Additionally, the results of this study suggest that stereoscopic visualization can be differently effective depending on the type of technology used. More importantly, the findings suggest that other possible mechanisms are responsible for the acquired 3D effect and positive effect on learning. Next research should focus on the working mechanisms that explain the effectiveness of stereoscopic 3DVT. Only by knowing why particular 3D technology works will enable educators and researcher to properly design and implement this tool in medical education.

CONCLUSIONS

In summary, binocular disparity alone does not contribute to better learning of anatomy in a stereoscopic 3D AR environment. Motion parallax, enabled by dynamic exploration, should be considered as a potential strong depth cue without or in combination with binocular disparity. Regardless of intervention, visual-spatial abilities were significantly and positively associated with the specimen test scores.

REFERENCES

1. 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.
2. Peterson DC, Mlynarczyk GS. Analysis of traditional versus three-dimensional augmented curriculum on anatomical learning outcome measures. Anat Sci Educ 2016;9:529–536.
3. 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.
4. Bergman EM, van der Vleuten CP, Scherpbier AJ. Why don't they know enough about anatomy? A narrative review. Med Teach 2011;33:403–409.
5. 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.
6. 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 2003;37:954–961.
7. Spielmann PM, Oliver CW. The carpal bones: A basic test of medical students and junior doctors' knowledge of anatomy. Surgeon 2005;3:257–259.
8. 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.
9. Bergman EM, Prince KJ, Drukker J, van der Vleuten CP, Scherpbier AJ. How much anatomy is enough? Anat Sci Educ 2008;1:184–188.
10. Cook DA. The research we still are not doing: An agenda for the study of computer-based learning. Acad Med 2005;80:541–548.
11. Cutting JE, Vishton PM. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In: Epstein W, Rogers S (Editors). Perception of Space and Motion. 1st Ed. San Diego, CA: Academic Press. 1995. p 69–117.
12. Johnston EB, Cumming BG, Landy MS. Integration of stereopsis and motion shape cues. Vision Res 1994;34:2259–2275.
13. 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.
14. Kytö M, Mäkinen A, Tossavainen T, Oittinen PT.Stereoscopic depth perception in video see-through augmented reality within action space. J Electron Imag 2014;23:011006.
15. Martinez EM, Junke B, Holub J, Hisley K, Eliot D, Winer E. Evaluation of monoscopic and stereoscopic displays for visual-spatial tasks in medical contexts. Comput Biol Med 2015;61:138–143.
16. Railo H, Saastamoinen J, Kylmala S, Peltola A. Binocular disparity can augment the capacity of vision without affecting subjective experience of depth. Sci Rep 2018;8:15798.
17. Anderson SJ, Jamniczky HA, Krigolson OE, Coderre SP, Hecker KG. Quantifying two-dimensional and three-dimensional stereoscopic learning in anatomy using electroencephalography. NPJ Sci Learn 2019;4:10.
18. Garg A, Norman GR, Spero L, Maheshwari P. Do virtual computer models hinder anatomy learning? Acad Med 1999;74:S87–S89.
19. Garg A, Norman G, Spero L, Taylor I. Learning anatomy: Do new computer models improve spatial understanding? Med Teach 1999;21:519–522.

20. 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.
21. 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.
22. 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.
23. Bogomolova K, van der Ham IJM, Dankbaar MEW, van den Broek WW, Hovius SER, van der Hage JA, Hierck BP. The effect of stereoscopic augmented reality visualization on learning anatomy and the modifying effect of visual-spatial abilities: a double-center randomized controlled trial. Anat Sci Educ 2020;13:558–567.
24. Hegarty M, Sims VK. Individual differences in use of diagrams as external memory in mechanical reasoning. Mem Cognit 1994;22:411–430.
25. Mayer RE. Multimedia Learning. 3rd Ed. Cambridge, UK: Cambridge University Press. 2001. 452 p.
26. Huk T. Who benefits from learning with 3D models? The case of spatial ability. J Comput Assist Learn 2006;22:392–404.
27. 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.
28. 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.
29. Kang X, Azizian M, Wilson E, Wu K, Martin AD, Kane TD, Peters CA, Cleary K, Shekhar R. Stereoscopic augmented reality for laparoscopic surgery. Surg Endosc 2014;28:2227–2235.
30. Cutolo F, Freschi C, Mascioli S, Parchi PD, Ferrari M, Ferrari V. Robust and accurate algorithm for wearable stereoscopic augmented reality with three indistinguishable markers. Electronics 2016;5:59.
31. Cutolo F, Meola A, Carbone M, Sinceri S, Cagnazzo F, Denaro E, Esposito N, Ferrari M, Ferrari V. A new head-mounted display-based augmented reality system in neurosurgical oncology: A study on phantom. Comput Assist Surg 2017;22:39–53.
32. Sommer B, Barnes DG, Boyd S, Chandler T, Cordeil M, Czauderna T, Klapperstück M, Klein K, Nguyen TD, Nim H, Stephens K. 3D-stereoscopic immersive analytics projects at Monash University and University of Konstanz. Electron Imag 2017:179–187.
33. Birt J, Stromberga Z, Cowling M, Moro C. Mobile mixed reality for experiential learning and simulation in medical and health sciences education. Information 2018;9:31.
34. Checcucci E, Amparore D, Pecoraro A, Peretti D, Aimar R, De Cillis S, Piramide F, Volpi G, Piazzolla P, Manfrin D, Manfredi M, Fiori C, Porpiglia F. 3D mixed reality holograms for preoperative surgical planning of nephron-sparing surgery: evaluation of surgeons' perception. Minerva Urol Nefrol. 2021;73:367-375.
35. Kumar RP, Pelanis E, Bugge R, Bruna H, Palomara R, Aghayan DL, Fretland AA, Edwin B, Elle OJ. Use of mixed reality for surgery planning: Assessment and development workflow. J Biomed Inform 2017;8:1000777
36. Schwab K, Smith R, Brown V, Whyte M, Jourdan I. Evolution of stereoscopic imaging in surgery and recent advances. World journal of gastrointestinal endoscopy 2017;9:368–377.

37. Itatani Y, Obama K, Nishigori T, Ganeko R, Tsunoda S, Hosogi H, Hisamori S, Hashimoto K, Sakai Y. Three-dimensional Stereoscopic Visualization Shortens Operative Time in Laparoscopic Gastrectomy for Gastric Cancer. Sci Rep 2019;9: 4108
38. Bogomolova K, Hierck BP, Looijen AEM, Pilon JNM, Putter H, Wainman B, Hovius SER, van der Hage JA. Stereoscopic three-dimensional visualisation technology in anatomy learning: A meta-analysis. Med Educ 2021;55(3):317-327.
39. Moro C, Phelps C, Redmond P, Stromberga Z. HoloLens and mobile augmented reality in medical and health science education: A randomised controlled trial. Br J Educ Technol 2021;52:680-694.
40. Wainman B, Pukas G, Wolak L, Mohanraj S, Lamb J, Norman GR. The critical role of stereopsis in virtual and mixed reality learning environments. Anat Sci Educ 2020;13:398–405
41. Cook DA, Beckman TJ. Reflections on experimental research in medical education. Adv Health Sci Educ Theory Pract 2010;15:455–464.
42. Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. Percept Mot Skills 1978;47:599–604.
43. 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.
44. 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.
45. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities and anatomy knowledge assessment: A systematic review. Anat Sci Educ 2017;10:235–241.
46. Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR. Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain. 1956. 1st Ed. New York, NY: David McKay Company. 207 p.
47. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Adv Psychol 1988;52:139–83.
48. Rogers B, Graham M. Motion parallax as an independent cue for depth-perception. Perception 1979;8:125–134.
49. Bradshaw MF, Rogers BJ. The interaction of binocular disparity and motion parallax in the computation of depth. Vision Res 1996;36:3457–3468.
50. Naepflin U, Menozzi M. Can movement parallax compensate lacking stereopsis in spatial explorative search tasks? Displays 2001;22:157–164.
51. Aygar E, Ware C, Rogers D. The contribution of stereoscopic and motion depth cues to the perception of structures in 3D point clouds. ACM Trans Appl Percept 2018;15:9.
52. Lankheet MJ, Palmen M. Stereoscopic segregation of transparent surfaces and the effect of motion contrast. Vision Res 1998;38:659–668.
53. Sherman A, Papathomas TV, Jain A, Keane BP. The role of stereopsis, motion parallax, perspective and angle polarity in perceiving 3-D shape. Seeing Perceiving 2012;25:263–285.
54. Oh CS, Won HS, Kim KJ, Jang DS. "Digit anatomy": A new technique for learning anatomy using motor memory. Anat Sci Educ 2011;4:132–141.
55. Cherdieu M, Palombi O, Gerber S, Troccaz J and Rochet-Capellan A. Make Gestures to Learn: Reproducing Gestures Improves the Learning of Anatomical Knowledge More than Just Seeing Gestures. Psychol 2017:1689.
56. Dickson AD and Stephens BW. It's all in the mime: actions speak louder than words when teaching the cranial nerves. Anat Sci Educ 2015;8:584-592.

57. Roach VA, Mi M, Mussell J, Van Nuland SE, Lufler RS, DeVeau KM, Dunham SM, Husmann P, Herriott HL, Edwards DN, Doubleday AF, Wilson BM, Wilson AB. Correlating spatial ability with anatomy assessment performance: A meta-analysis. Anat Sci Educ 2021;14:317-329.
58. Luursema JM, Verwey WB, Kommers PA, Geelkerken RH, Vos HJ. Optimizing conditions for computer-assisted anatomical learning. Interact Comput 2006;18:1123–1138.
59. Hegarty M, Burte H, Boone AP. Individual differences in large-scale spatial abilities and strategies In: Montello DR (Editor). Handbook of Behavioral and Cognitive Geography. 2018. 1st Ed. Cheltenham, UK: Edward Elgar Publishing Limited; p 231–246.
60. Baenninger M, Newcombe N. The role of experience in spatial test performance: A meta-analysis. Sex Roles 1989;20:327–344.
61. Langlois J, Wells GA, Lecourtois M, Bergeron G, Yetisir E, Martin M. Sex differences in spatial abilities of medical graduates entering residency programs. Anat Sci Educ 2013;6:368–375.
62. Uttal DH, Meadow NG, Tipton E, Hand LL, Alden AR, Warren C, Newcombe NS. The malleability of spatial skills: A meta-analysis of training studies. Psychol Bull 2013;139:352–402.
63. Nguyen N, Mulla A, Nelson AJ, Wilson TD. Visuospatial anatomy comprehension: The role of spatial visualization and problem-solving strategies. Anat Sci Educ 2014;7:280–288.
64. Guimarães B, Firmino-Machado J, Tsisar S, Viana B, Pinto-Sousa M, Vieira-Marques P, Cruz-Correia R, Ferreira MA. 2019. The role of anatomy computer-assisted learning on spatial abilities of medical students. Anat Sci Educ 12:138–153.
65. Vorstenbosch MA, Klaassen TP, Donders AR, Kooloos JG, Bolhuis SM, Laan RF. Learning anatomy enhances spatial ability. Anat Sci Educ 2013;6:257–262..