

# 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,<sup>1</sup>  <sup>1,2\*</sup> Ineke J.M. van der Ham,<sup>3</sup> Mary E.W. Dankbaar,<sup>4</sup> Walter W. van den Broek,<sup>4</sup> Steven E.R. Hovius,<sup>5,6</sup> Jos A. van der Hage,<sup>1,2</sup> Beerend P. Hierck<sup>2,7,8,9</sup>

<sup>1</sup>Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands

<sup>2</sup>Center for Innovation of Medical Education, Leiden University Medical Center, Leiden, The Netherlands

<sup>3</sup>Department of Neuropsychology, Leiden University, Leiden, The Netherlands

<sup>4</sup>Institute for Medical Education Research Rotterdam, Rotterdam Erasmus University Medical Center, Rotterdam, The Netherlands

<sup>5</sup>Department of Plastic and Reconstructive Surgery and Hand Surgery, Rotterdam Erasmus University Medical Center, Rotterdam, The Netherlands

<sup>6</sup>Department of Plastic and Reconstructive Surgery, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>7</sup>Department of Anatomy and Embryology, Leiden University Medical Center, Leiden, The Netherlands

<sup>8</sup>Centre for Innovation, Leiden University, The Hague, The Netherlands

<sup>9</sup>Leiden Teachers' Academy, Leiden University, Leiden, The Netherlands

Monoscopically projected three-dimensional (3D) visualization technology may have significant disadvantages for students with lower visual-spatial abilities 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 visual-spatial abilities on learning. In a double-center randomized controlled trial, first- and second-year (bio)medical undergraduates studied lower limb anatomy with stereoscopic 3D AR model ( $n = 20$ ), monoscopic 3D desktop model ( $n = 20$ ), or two-dimensional (2D) anatomical atlas ( $n = 18$ ). Visual-spatial abilities were tested with Mental Rotation Test (MRT), Paper Folding Test (PFT), and Mechanical Reasoning (MR) Test. Anatomical knowledge was assessed by the validated 30-item paper posttest. The overall posttest scores in the stereoscopic 3D AR group (47.8%) were similar to those in the monoscopic 3D desktop group (38.5%;  $P = 0.240$ ) and the 2D anatomical atlas group (50.9%;  $P = 1.00$ ). When stratified by visual-spatial abilities test scores, students with lower MRT scores achieved higher posttest scores in the stereoscopic 3D AR group (49.2%) as compared to the monoscopic 3D desktop group (33.4%;  $P = 0.015$ ) and similar to the scores in the 2D group (46.4%;  $P = 0.99$ ). Participants with higher MRT scores performed equally well in all conditions. It is instrumental to consider an aptitude–treatment interaction caused by visual-spatial abilities 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. *Anat Sci Educ* 13: 558–567. © 2019 The Authors. Anatomical Sciences Education published by Wiley Periodicals, Inc. on behalf of American Association of Anatomists.

**Key words:** gross anatomy education; stereoscopic three-dimensional visualization technology; visual-spatial abilities; mental rotation test; undergraduate education; augmented reality

\*Correspondence to: Dr. Katerina Bogomolova, Department of Surgery, Leiden University Medical Center, PO Box 9600, Zone D6-P, 2300 RC Leiden, The Netherlands. E-mail: [k.bogomolova@lumc.nl](mailto:k.bogomolova@lumc.nl)

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

Received 26 June 2019; Revised 20 December 2019; Accepted 27 December 2019.

Published online 27 January 2020 in Wiley Online Library ([wileyonlinelibrary.com](http://wileyonlinelibrary.com)). DOI 10.1002/ase.1941

© 2019 The Authors. Anatomical Sciences Education published by Wiley Periodicals, Inc. on behalf of American Association of Anatomists

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## INTRODUCTION

Anatomical knowledge among undergraduate medical students and recently graduated doctors has repeatedly been reported to be insufficient (McKeown et al., 2003; Prince et al., 2005; Spielmann and Oliver, 2005; Waterston and Stewart, 2005; Bergman et al., 2008). One of the main reasons is the decrease in anatomy teaching time in undergraduate education, related to increasing costs and limited availability of cadavers, and time pressure on the curriculum, have led to a decreased exposure to traditional cadaveric dissections (Pryde and Black, 2005; Waterston and Stewart, 2005; Azer and Eizenberg, 2007; Drake et al., 2009; Bergman et al., 2013). Although, the educational value is being debated (Azer and Eizenberg, 2007), 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 (Klatzky and Lederman, 2011; Reid et al., 2018).

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 (2015) 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 lower visual-spatial abilities (Garg et al., 1999a,b; 2002; Levinson et al., 2007; Naaz, 2012). The disadvantages are well known in the research field of 3D learning and were first described by Garg and colleagues (1999a,b; 2002). In these studies, visual-spatial abilities 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 (Garg et al., 2002; Huk, 2006; Levinson et al., 2007; Khot et al., 2013).

However, when traditional digital 3D models are viewed stereoscopically by projecting a slightly shifted image to the left and right eye, the disadvantages for students with lower visual-spatial abilities seem to disappear. Cui and colleagues (2017) 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. They reported a better performance of undergraduate medical students after learning anatomy with a stereoscopic 3D model. Most importantly, students with lower visual-spatial abilities have improved their knowledge test scores to a level comparable to that demonstrated by the students with higher visual-spatial abilities. The role of stereopsis has also been evaluated by Luursema and colleagues (2006, 2008, 2017) 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 visual-spatial abilities and the stereoscopic condition remained significant (Luursema et al., 2008). Overall, stereoscopic 3DVT appears to have a positive effect on learning as recently

demonstrated by Hackett and Proctor (2018). 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 visual-spatial abilities 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 (2018). 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. First, 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 (Jolicoeur and Milliken, 1989; Kourtzi et al., 2003; Luursema et al., 2006; Verhoef et al., 2016). Stereopsis might therefore facilitate a better comprehension of anatomy especially among students with lower visual-spatial abilities. Second, the reported differences in learning effect between students with lower and higher levels of visual-spatial abilities 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 (Cook, 2005). 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  $\times$  intervention), as demonstrated by Luursema and colleagues (2008) and Cui and colleagues (2017).

## Augmented Reality in Anatomy 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 (Moro et al., 2017; Kuehn 2018). 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 (HMDs) such as AR and VR devices. Whether the anatomy can be perceived in a real 3D plane, depends on the type of device. From flat screens visualization of 3D content is usually obtained monoscopically with various interactive features added to the digital overlay provided by these devices (Küçük et al., 2016; Barmaki et al., 2019; Sugiura et al., 2019). HMD can provide an interactive and stereoscopic way of 3D visualization (Supplementary Material 1). With AR technology, such as with the HoloLens<sup>®</sup>, the most distinguishing feature is the ability to perceive an anatomical model in a real 3D plane without losing the 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 (McJunkin et al., 2018). The educational effectiveness of this technology for teaching anatomy has not been evaluated yet. For the purpose of this study an augmented reality 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.

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.

## 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 visual-spatial abilities would modify the observed learning effect. In addition, 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 visual-spatial abilities benefit most from the stereoscopic 3D view of the model.

## 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 posttest (Cook and Beckman, 2010). 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 15 euros at the completion of the experimental session.

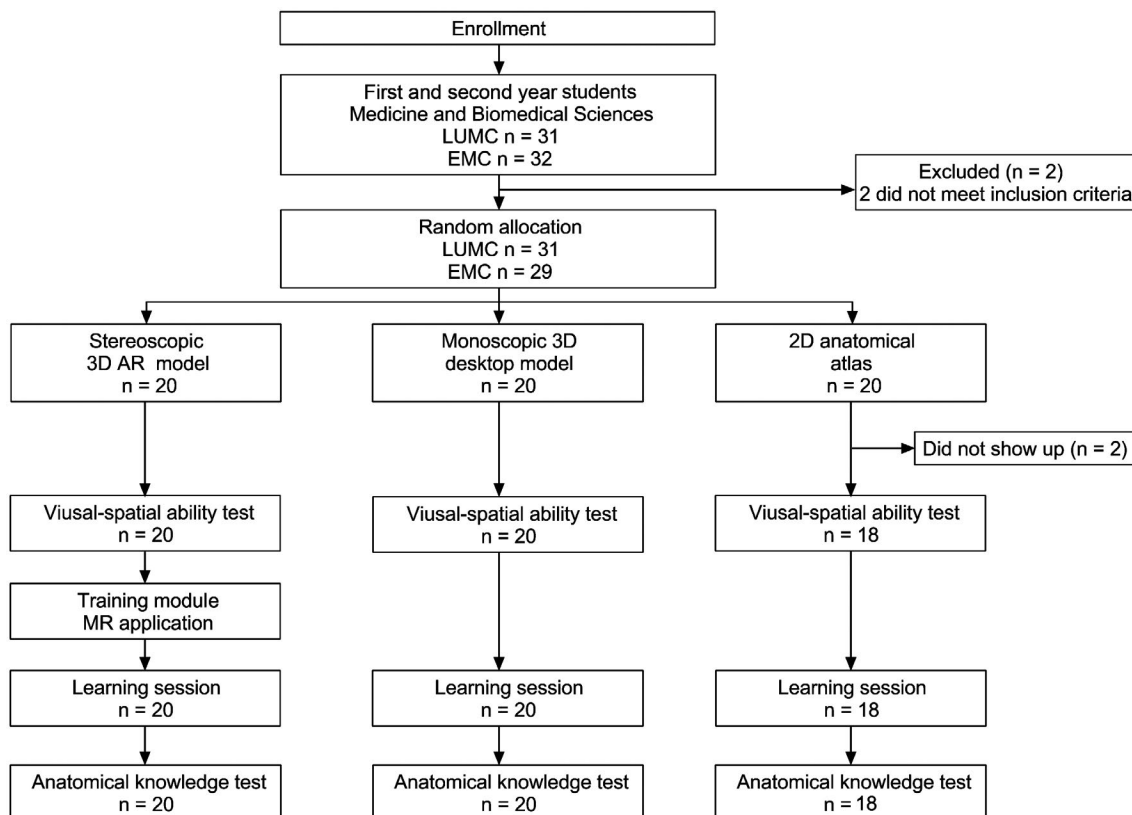
## MATERIALS AND METHODS

### Study Design

A double-center randomized controlled trial was conducted at the Leiden University Medical Center (LUMC) and the Erasmus

### 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



**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.

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 augmented reality application *DynamicAnatomy* (LUMC, 2019) for HoloLens<sup>®</sup>, version 1.0 (Microsoft Corp., Redmond, WA) 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](http://www.mr4education.com)). The application represented a dynamic and fully interactive stereoscopic 3D model of the lower leg. The model was presented as a 3D virtual object in the physical space (Supplementary Material 1A). The HoloLens<sup>®</sup> glasses are transparent which enabled participants to stereoscopically interact with the model without losing the sense of their physical environment. A unique feature includes 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 (Table 1). 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-minute 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 1B). All other features such as voice control, auditory feedback, and scaling were unchanged (Table 1).

In the 2D anatomical atlas group, study material included selected handouts from an anatomy atlas (Putz and Pabst, 2006) and an anatomy textbook (Moore et al., 2013) 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 (Biggs, 1996) (Supplementary Material 2A, 2B). The learning goals were formulated and organized according to Bloom's Taxonomy of Learning Objectives (Bloom et al., 1956). 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.

## Visual-Spatial Abilities Assessment

Visual-spatial abilities 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) (Shepard and Metzler, 1971), previously validated by Vandenberg and Kuse (1978) and redrawn by Peters and colleagues (1995) (Supplementary Material 3A). This psychometric test is being widely used in the assessment of visual-spatial abilities and has repeatedly shown its positive association with anatomy learning and assessment (Guillot et al., 2007;

**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  | +                        | +                           | -                   |

Positive sign (+), a feature is present; negative sign (-), a feature is not present; AR, augmented reality; 3D, three-dimensional; 2D, two-dimensional.

Langlois et al., 2017). 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 visual-spatial abilities, were measured by the 10-item Paper Folding Test (PFT), previously validated by Ekstrom and colleagues (1976) (Supplementary Material 3B). The post-hoc level of internal consistency (Cronbach's alpha) of the PFT in this study was 0.76. In addition, mechanical reasoning was measured by a standardized 12-item Mechanical Reasoning (MR) Test developed for this experiment at the Department of Neuropsychology (Supplementary Material 3C). The post-hoc level of internal consistency (Cronbach's alpha) of the MR test in this study was 0.76. Duration of the assessment was three minutes for each test. After 3 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 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 4). Content validation was performed by two experts in the field of anatomy and plastic and reconstructive surgery. The test was then piloted among 12 medical students for item clarity. The post-hoc calculated level of internal consistency (Cronbach's alpha) was 0.78. 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 5-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  $\chi^2$  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  $\times k$ ). The results were stratified by MRT, PFT, and MR test scores to evaluate possible aptitude-treatment interaction between visual-spatial abilities and type of intervention. In addition, an 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, visual-spatial abilities test score as a covariate, and "visual-spatial abilities test score"  $\times$  "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 (Cohen, 1988). 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  $P < 0.05$ .

## RESULTS

A total of 60 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 baseline characteristics of the 58 participants.

### 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%,  $SD \pm 9.8$ ) performed equally well on the knowledge test as the monoscopic 3D desktop group (38.5%,  $SD \pm 14.3$ ;  $F(2,54) = 4.79$ ;  $P = 0.240$ ) and the 2D anatomical atlas group (50.9%,  $SD \pm 13.8$ ;  $F(2,54) = 4.79$ ;  $P = 1.00$ ). The 2D anatomical atlas group, however, outperformed the monoscopic 3D desktop group ( $F(2,54) = 4.79$ ;  $P = 0.042$ ).

### Scores Stratified by Visual-Spatial Abilities

When total scores on the anatomy knowledge test were stratified by MRT, PFT, and 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 ( $F(2,23) = 0.83$ ,  $P = 0.448$ ). However, among MRT-low participants significant differences were found between groups. The stereoscopic 3D AR group (49.2%,  $SD \pm 9.5$ ) significantly outperformed the monoscopic 3D desktop group (33.4%,  $SD \pm 11.5$ ;  $F(2,28) = 6.59$ ,  $P = 0.015$ , Cohen's  $d = 1.54$ ), and performed equally well as the 2D anatomical atlas group (46.4%,  $SD \pm 14.5$ ;  $F(2,28) = 6.59$ ,  $P = 0.990$ , 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 ( $P = 0.250$ ). The MRT-low group performed significantly worse than the MRT-high students in the monoscopic 3D desktop group (33.4%,  $SD \pm 11.5$  vs. 49.7%,  $SD \pm 13.9$ ;  $P = 0.015$ , Cohen's  $d = -1.3$ ) However, they performed equally well in the stereoscopic 3D AR and 2D anatomical atlas groups.

The observed differences strongly indicate an aptitude-treatment effect caused by visual-spatial abilities. This

**Table 2.**

Baseline Characteristics of the Included Participants

| Characteristic                                   | Stereoscopic 3D AR model; n = 20 | Monoscopic 3D desktop model; n = 20 | 2D anatomical atlas; n = 18 | P-value |
|--|----------------------------------|-------------------------------------|-----------------------------|---------|
| Sex, n (%)                                       |                                  |                                     |                             |         |
| Male   | 8 (40.0)                         | 6 (35.0)                            | 7 (39.0)                    | 0.773   |
| Female   | 12 (60.0)                        | 13 (65.0)                           | 11 (61.0)                   |         |
| Age, mean ( $\pm$ SD), years                     | 18.5 (0.8)                       | 18.7 (1.0)                          | 18.7 (0.8)                  | 0.720   |
| Medical Center, n (%)                            |                                  |                                     |                             |         |
| Leiden University MC                             | 10 (50.0)                        | 11 (55.0)                           | 10 (55.6)                   | 0.929   |
| 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)                   | 0.842   |
| Biomedical sciences                              | 3 (15.0)                         | 4 (20.0)                            | 4 (22.2)                    |         |
| Study year, n (%)                                |                                  |                                     |                             |         |
| First year                                       | 17 (85.0)                        | 18 (90.0)                           | 16 (88.9)                   | 0.879   |
| Second year                                      | 3 (15.0)                         | 2 (10.0)                            | 2 (11.1)                    |         |
| Visual-spatial abilities score, mean ( $\pm$ SD) |                                  |                                     |                             |         |
| Mental Rotation Test                             | 7.1 (2.9)                        | 6.0 (2.4)                           | 8.4 (2.1)                   | 0.090   |
| Paper Folding Test                               | 6.2 (1.8)                        | 6.5 (2.6)                           | 7.6 (2.2)                   | 0.104   |
| Mechanical Reasoning Test                        | 9.3 (2.6)                        | 9.3 (2.1)                           | 9.3 (3.2)                   | 0.990   |

P-value < 0.05 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.

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”  $\times$  “intervention group” showed a marginal trend toward significance ( $F(2) = 3.04$ ;  $P = 0.05$ ). 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 ( $4.8 \pm 0.4$  vs.  $3.4 \pm 0.8$  vs.  $2.4 \pm 0.9$ ;  $F(2,54) = 50.3$ ,  $P = 0.003$ ). 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 ( $4.3 \pm 0.6$  vs.  $4.1 \pm 0.9$  vs.  $4.1 \pm 0.8$ ;  $F(2,54) = 0.6$ ,  $P = 0.574$ ).

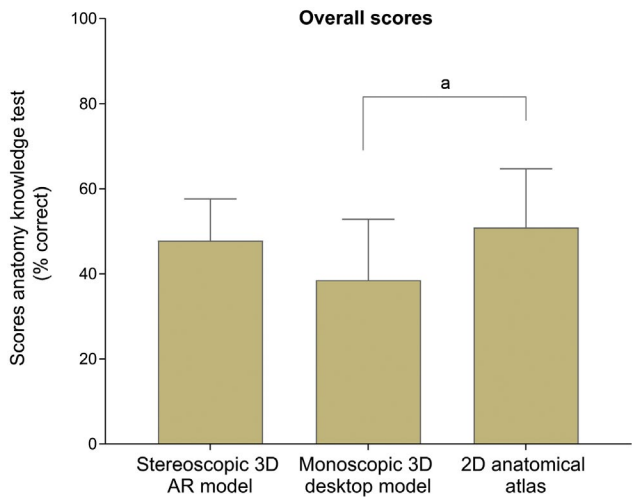
## DISCUSSION

This study aimed to investigate the educational effectiveness of learning with stereoscopic AR visualization technology and

to evaluate whether visual-spatial abilities would modify the learning effect.

First, the observed aptitude–treatment interaction caused by visual-spatial abilities 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 visual-spatial abilities scores as measured by the MRT. These differences were detectable only after stratification of the overall results pointing toward an aptitude–treatment interaction, also referred to as “effect measure modification” (Cook, 2005; Rothman et al., 2008; Corraini et al., 2017). 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 visual-spatial abilities were treated only as a confounder, in the absence of stratification, the differences between monoscopic and stereoscopic conditions for different levels of visual-spatial abilities 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 covariate in the regression analysis), will still not be sufficient, and the results can still be misleading.

Second, 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



**Figure 2.**

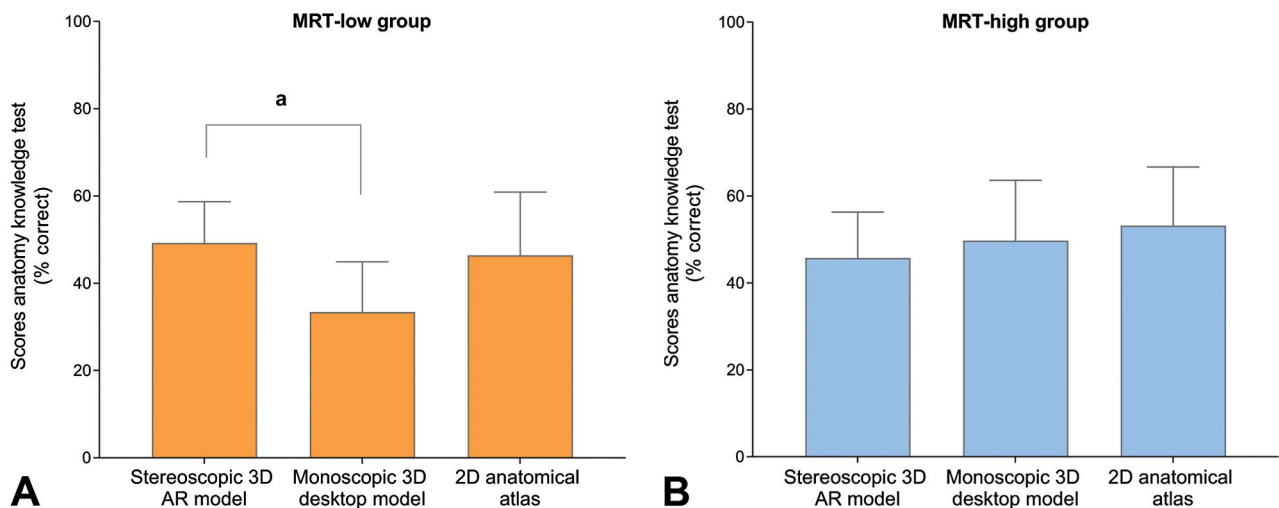
Differences in overall mean percentages correct answers on the anatomy knowledge test between three educational interventions. <sup>a</sup> $P < 0.05$  analysis of variance with a Bonferroni correction for multiple comparison. MRT, Mental Rotation Test; AR, augmented reality; 3D, three-dimensional; 2D, two-dimensional.

disadvantages for students with lower visual-spatial abilities (Garg et al., 1999a,b; 2002; Levinson et al., 2007; Naaz, 2012). It has been hypothesized that 3D objects are memorized as key view-based 2D images (Bulthoff et al., 1995; Garg et al., 2002). Viewing an unfamiliar 3D object from multiple angles, could therefore lead to an increase in extraneous cognitive load (Huk, 2006; Khot et al., 2013; Mayer, 2014). 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 (Jolicoeur and

Milliken, 1989; Kourtzi et al., 2003; Luursema et al., 2006). 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 (Verhoef et al., 2016). 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 visual-spatial abilities.

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. Moreover, 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 (Cook 2005). This eliminates all possible confounding effects of additional features such as hand gestures, that can vary between different types of media.

Third, 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



**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 ( $n = 26$ ). <sup>a</sup> $P < 0.05$  analysis of variance with a Bonferroni correction for multiple comparison. MRT, Mental Rotation Test; AR, augmented reality; 3D, three-dimensional; 2D, two-dimensional.

**Table 3.**

Differences in Students' Learning Experience between Three Educational Interventions

| Statement  | Stereoscopic 3D AR model; n = 20<br>mean ( $\pm$ SD) | Monoscopic 3D desktop model; n = 20<br>mean ( $\pm$ SD) | 2D anatomical atlas; n = 18<br>mean ( $\pm$ SD) | P-value            |
|--|--|---|---|--------------------|
| The study time was long enough to study the required number of anatomical structures | 3.0 ( $\pm$ 0.9)                                     | 2.3 ( $\pm$ 0.7)  | 2.6 ( $\pm$ 0.9)                                | 0.192              |
| The questions in anatomy test were representative for the studied material           | 3.8 ( $\pm$ 0.6)                                     | 3.8 ( $\pm$ 0.6)  | 3.7 ( $\pm$ 0.7)                                | 0.709              |
| I enjoyed studying with ...  | 4.8 ( $\pm$ 0.4) <sup>b</sup>                        | 3.4 ( $\pm$ 0.8) <sup>b</sup>                           | 2.4 ( $\pm$ 0.9) <sup>b</sup>                   | 0.003 <sup>a</sup> |
| Learning material was easy to use  | 4.3 ( $\pm$ 0.6) <sup>c</sup>                        | 3.4 ( $\pm$ 0.9) <sup>c</sup>                           | 3.0 ( $\pm$ 0.8) <sup>c</sup>                   | 0.009 <sup>a</sup> |
| My knowledge about anatomy of the lower leg is improved after studying with ...      | 4.3 ( $\pm$ 0.6)                                     | 4.1 ( $\pm$ 0.9)  | 4.1 ( $\pm$ 0.8)                                | 0.574              |
| I would recommend studying with ... to my fellow students                            | 4.6 ( $\pm$ 0.5) <sup>b</sup>                        | 3.7 ( $\pm$ 0.8) <sup>b</sup>                           | 2.4 ( $\pm$ 0.9) <sup>b</sup>                   | 0.003 <sup>a</sup> |

Response options on a five-point Likert scale ranged from 1 = very dissatisfied to 5 = very satisfied. Scores are expressed in means ( $\pm$ SD). <sup>a</sup> $P < 0.05$  analysis of variance with a Bonferroni correction for multiple comparison; <sup>b</sup>significant difference between all the three groups; <sup>c</sup>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.

technology versus the use of prosected cadaveric specimens, students have performed best on the identification questions aligned with the respective study materials (Mitrousias et al., 2018, 2020). A similar effect has been reported by Henssen and colleagues (2020) 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 (2010) 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. 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 (Smallman and Hegarty, 2007). In the current study, relevant 2D images were selected from 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). In addition, students with lower visual-spatial abilities 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 (Stull et al., 2009; Hegarty, 2010). 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 (2017a, b; 2019) in performing a mental rotation task. Students with higher visual-spatial abilities had a distinct eye movement pattern in solving mental rotation tasks than low performing students (Roach et al., 2017a). 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 (Roach et al., 2019). 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 visual-spatial abilities 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 in order 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 (McKeown et al., 2003; Prince et al., 2005; Spielmann and Oliver, 2005; Waterston and Stewart, 2005; Bergman et al., 2008). When designing new virtual reality (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.

## Limitation of the Study

There are some methodological limitations in this study. First, due to the limited availability of hardware, the study was restricted to a maximum of 20 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. Third, the participants were not tested for their lack of depth perception which could be present in about 5% of the study population (Mather, 2006). Based on these statistics, 1-2 of the 20 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. In addition, 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 in students with lower visual-spatial abilities. 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.

## ACKNOWLEDGMENTS

The authors sincerely thank Prof. Dr. Marco C. de Rooter for content validation of the anatomical knowledge test and Renée A. Hendriks, M.Sc. for the verification of the alignment construct of the learning objectives and assessment. The authors have no conflicts of interest to disclose.

## NOTES ON CONTRIBUTORS

KATERINA BOGOMOLOVA, M.D., is a graduate (Ph.D.) student in the Department of Surgery and Center for Innovation in Medical Education, Leiden University Medical Center, in Leiden, The Netherlands. She is investigating the role of three-dimensional visualization technologies in anatomical and surgical education in relation to learners' spatial abilities.

INEKE J.M. VAN DER HAM, Ph.D., is an associate professor neuropsychology in the Department of Health, Medical and Neuropsychology at Leiden University, in Leiden, The Netherlands. She focuses her research on spatial cognition in healthy and neuropsychological populations and on methodological considerations in the use of virtual reality.

MARY E.W. DANKBAAR, Ph.D., is a program manager of e-learning and assistant professor at Erasmus University Medical Center Rotterdam in Rotterdam, The Netherlands. She has a background in educational science and has designed and implemented several blended educational programs for the medical curriculum. Her focus in research is on designing simulation and games for skills training.

WALTER W. VAN DEN BROEK, M.D., Ph.D., is the director of medical education at the Erasmus University Medical Center Rotterdam and the scientific director of the Institute for Medical Education Research Rotterdam in Rotterdam, The Netherlands. He is also director of residency training in psychiatry at the Erasmus MC, teaches psychiatry to medical students, and is a coach for bachelor medical students.

STEVEN E.R. HOVIUS, M.D., Ph.D., is an emeritus professor of plastic and reconstructive surgery and hand surgery at Erasmus University Medical Center in Rotterdam and Radboud University Medical Center in Nijmegen, The Netherlands. He teaches anatomy of upper and lower limbs to medical students and surgical residents.

JOS A. VAN DER HAGE, M.D., Ph.D., is a professor of intra-curricular education in surgery in the Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands. He teaches surgery and anatomy to residents and medical students. One of his educational research topics is on 3D learning in anatomical and surgical education.

BEEREND P. HIERCK, Ph.D., is an assistant professor of anatomy in the Department of Anatomy and Embryology and researcher at the Center for Innovation in Medical Education, Leiden University Medical Center in Leiden, The Netherlands. He teaches anatomy, developmental biology and histology to (bio)medical students. His educational research focuses on 3D learning in the (bio)medical curriculum, with a special interest in the use of extended reality.

## LITERATURE CITED

- Azer SA, Eizenberg N. 2007. Do we need dissection in an integrated problem-based learning medical course? Perceptions of first- and second-year students. *Surg Radiol Anat* 29:173-180.
- Barmaki R, Yu K, Pearlman R, Shingles R, Bork F, Osgood GM, Navab N. 2019. Enhancement of anatomical education using augmented reality: An empirical study of body painting. *Anat Sci Educ* 12:599-609.
- Bergman EM, Prince KJ, Drukker J, van der Vleuten CP, Scherpbier AJ. 2008. How much anatomy is enough? *Anat Sci Educ* 1:184-188.
- Bergman EM, de Bruin AB, Herrler A, Verhrijen IW, Scherpbier AJ, van der Vleuten CP. 2013. Students' perceptions of anatomy across the undergraduate problem-based learning medical curriculum: A phenomenographical study. *BMC Med Educ* 13:152-162.

- Biggs J. 1996. Enhancing teaching through constructive alignment. *High Educ* 32:347–364.
- Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR. 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. 1st Ed. New York, NY: David McKay Company. 207 p.
- Bulthoff HH, Edelman SY, Tarr MJ. 1995. How are three-dimensional objects represented in the brain? *Cereb Cortex* 5:247–260.
- Cohen J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Ed. Hillsdale, NJ: Lawrence Erlbaum Associates. 400 p.
- Cook DA. 2005. The research we still are not doing: An agenda for the study of computer-based learning. *Acad Med* 80:541–548.
- Cook DA, Beckman TJ. 2010. Reflections on experimental research in medical education. *Adv Health Sci Educ Theory Pract* 15:455–464.
- Corraini P, Olsen M, Pedersen L, Dekkers OM, Vandenbroucke JP. 2017. Effect modification, interaction and mediation: An overview of theoretical insights for clinical investigators. *Clin Epidemiol* 9:331–338.
- Cui D, Wilson TD, Rockhold RW, Lehman MN, Lynch JC. 2017. Evaluation of the effectiveness of 3D vascular stereoscopic models in anatomy instruction for first year medical students. *Anat Sci Educ* 10:34–45.
- Drake RL, McBride JM, Lachman N, Pawlina W. 2009. Medical education in the anatomical sciences: The winds of change continue to blow. *Anat Sci Educ* 2:253–259.
- Ekstrom RB, French J, Harman HH, Dermen D. 1976. *Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service. 314 p.
- Garg A, Norman GR, Spero L, Maheshwari P. 1999a. Do virtual computer models hinder anatomy learning? *Acad Med* 74:S87–S89.
- Garg A, Norman G, Spero L, Taylor I. 1999b. Learning anatomy: Do new computer models improve spatial understanding? *Med Teach* 21:519–522.
- Garg AX, Norman GR, Eva KW, Spero L, Sharan S. 2002. Is there any real virtue of virtual reality? The minor role of multiple orientations in learning anatomy from computers. *Acad Med* 77:S97–S99.
- Guillot A, Champely S, Batier C, Thiriet P, Collet C. 2007. Relationship between spatial abilities, mental rotation and functional anatomy learning. *Adv Health Sci Educ Theory Pract* 12:491–507.
- Hackett M, Proctor M. 2018. The effect of autostereoscopic holograms on anatomical knowledge: A randomized trial. *Med Educ* 52:1147–1155.
- Hegarty M. 2010. Chapter 7 - Components of spatial intelligence. *Psychol Learn Motiv* 52:265–297.
- 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* 13:353–365.
- Huk T. 2006. Who benefits from learning with 3D models? The case of spatial ability. *J Comput Assist Learn* 22:392–404.
- Jolicoeur P, Milliken B. 1989. Identification of disoriented objects: Effects of context of prior presentation. *J Exp Psychol Learn Mem Cogn* 15:200–210.
- Khot Z, Quinlan K, Norman GR, Wainman B. 2013. The relative effectiveness of computer-based and traditional resources for education in anatomy. *Anat Sci Educ* 6:211–215.
- Klatzky RL, Lederman SJ. 2011. Haptic object perception11Spatial dimensionality and relation to vision. *Philos Trans R Soc Lond B Biol Sci* 366:3097–3105.
- Kourtzi Z, Erb M, Grodd W, Bulthoff HH. 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cereb Cortex* 13:911–920.
- Küçük S, Kapakin S, Göktaş Y. 2016. Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. *Anat Sci Educ* 9:411–421.
- Kuehn BM. 2018. Virtual and augmented reality put a twist on medical education. *JAMA* 319:756–758.
- Langlois J, Bellemare C, Toulouse J, Wells GA. 2017. Spatial abilities and anatomy knowledge assessment: A systematic review. *Anat Sci Educ* 10:235–241.
- Levinson AJ, Weaver B, Garside S, McGinn H, Norman GR. 2007. Virtual reality and brain anatomy: A randomised trial of e-learning instructional designs. *Med Educ* 41:495–501.
- LUMC. 2019. Leiden University Medical Center. *DynamicAnatomy*. Application for HoloLens® requires Windows 10, version 10240.0 or higher, Xbox One. Centre for Innovation Leiden University, Leiden, The Netherlands. URL: <https://www.microsoft.com/en-us/p/dynamicanatomy/9nwlj4qq053p?SilentAuth=1&activetab=pivot:overviewtab> [accessed 10 December 2019].
- Luursema JM, Verwey WB, Kommers PA, Geelkerken RH, Vos HJ. 2006. Optimizing conditions for computer-assisted anatomical learning. *Interact Comput* 18:1123–1138.
- Luursema JM, Verwey WB, Kommers PA, Annema JH. 2008. The role of stereopsis in virtual anatomical learning. *Interact Comput* 20:455–460.
- Luursema JM, Vorstenbosch M, Kooloos J. 2017. Stereopsis, visuospatial ability, and virtual reality in anatomy learning. *Anat Res Int* 2017:1493135.
- Mather G. 2006. *Foundations of Perception*. 1st Ed. Hove, UK: Psychology Press. 400 p.
- Mayer RE. 2014. Cognitive theory of multimedia learning. In: Mayer RE (Editor). *Multimedia Learning*. 2nd Ed. Santa Barbara, CA: Cambridge University Press. p 43–47.
- McJunkin JL, Jiramongkolchai P, Chung W, Southworth M, Durakovic N, Buchman CA, Silva JR. 2018. Development of a mixed reality platform for lateral skull base anatomy. *Otol Neurotol* 39:e1137–e1142.
- McKeown PP, Heylings DJ, Stevenson M, McKelvey KJ, Nixon JR, McCluskey DR. 2003. The impact of curricular change on medical students' knowledge of anatomy. *Med Educ* 37:954–961.
- Mitrousias V, Karachalios TS, Varitimidis SE, Natsis K, Arvanitis DL, Zibis AH. 2020. Anatomy learning from prosected cadaveric specimens versus plastic models: A comparative study of upper limb anatomy. *Anat Sci Educ* 13:436–444.
- 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.
- Moore KL, Dalley AF, Agur AM. 2013. *Clinically Oriented Anatomy*. 7th Ed. Philadelphia, PA: Lippincott Williams & Wilkins. 1168 p.
- Moro C, Štromberga Z, Raikos A, Stirling A. 2017. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anat Sci Educ* 10:549–559.
- Naaz F. 2012. Learning from graphically integrated 2D and 3D representations improves retention of neuroanatomy. University of Louisville: Louisville, KY. Doctorate of Philosophy Dissertation. 76 p.
- Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C. 1995. A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain Cognit* 28:39–58.
- Prince KJ, Scherpbier AJ, Van Mameren H, Drukker J, van der Vleuten CP. 2005. Do students have sufficient knowledge of clinical anatomy? *Med Educ* 39:326–332.
- Pryde FR, Black SM. 2005. Anatomy in Scotland: 20 years of change. *Scott Med J* 50:96–98.
- Putz R, Pabst R. 2006. *Sobotta Atlas of Human Anatomy. Part 2*. 3rd Ed. Houten, The Netherlands: Bohn Stafleu van Loghum. 399 p.
- Reid S, Shapiro L, Louw G. 2018. How haptics and drawing enhance the learning of anatomy. *Anat Sci Educ* 12:164–172.
- Roach VA, Fraser GM, Kryklywy JH, Mitchell DGV, Wilson TD. 2017a. Different perspectives: Spatial ability influences where individuals look on a timed spatial test. *Anat Sci Educ* 10:224–234.
- Roach VA, Fraser GM, Kryklywy JH, Mitchell DG, Wilson TD. 2017b. Time limits in testing: An analysis of eye movements and visual attention in spatial problem solving. *Anat Sci Educ* 10:528–537.
- Roach VA, Fraser GM, Kryklywy JH, Mitchell DG, Wilson TD. 2019. Guiding low spatial ability individuals through visual cueing: The dual importance of where and when to look. *Anat Sci Educ* 12:32–42.
- Rothman KJ, Greenland S, Lash TL. 2008. *Modern Epidemiology*. 3rd Ed. Philadelphia, PA: Lippincott Williams and Wilkins. 758 p.
- Shepard RN, Metzler J. 1971. Mental rotation of three-dimensional objects. *Science* 171:701–703.
- Smallman HS, Hegarty M. 2007. 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.
- Spielmann PM, Oliver CW. 2005. The carpal bones: A basic test of medical students and junior doctors' knowledge of anatomy. *Surgeon* 3:257–259.
- Stull AT, Hegarty M, Mayer RE. 2009. Orientation references: Getting a handle on spatial learning. *J Educ Psychol* 101:803–816.
- Sugiura A, Kitama T, Toyoura M, Mao X. 2019. The use of augmented reality technology in medical specimen museum tours. *Anat Sci Educ* 12:561–571.
- Vandenberg SG, Kuse AR. 1978. Mental rotations, a group test of three-dimensional spatial visualization. *Percept Mot Skills* 47:599–604.
- Verhoef B-E, Vogels R, Janssen P. 2016. Binocular depth processing in the ventral visual pathway. *Philos Trans R Soc Lond B Biol Sci* 371:20150259.
- Wainman B, Wolak L, Pukas G, Zheng E, Norman GR. 2018. The superiority of three-dimensional physical models to two-dimensional computer presentations in anatomy learning. *Med Educ* 52:1138–1146.
- Waterston SW, Stewart IJ. 2005. Survey of clinicians' attitudes to the anatomical teaching and knowledge of medical students. *Clin Anat* 18:380–384.
- Yammine- K, Violato C. 2015. A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy. *Anat Sci Educ* 8:525–538.