



Universiteit
Leiden
The Netherlands

Anatomy dissection course improves the initially lower levels of visual-spatial abilities of medical undergraduates

Bogomolova, K.; Hierck, B.P.; Hage, J.A. van der; Hovius, S.E.R.

Citation

Bogomolova, K., Hierck, B. P., Hage, J. A. van der, & Hovius, S. E. R. (2019). Anatomy dissection course improves the initially lower levels of visual-spatial abilities of medical undergraduates. *Anatomical Sciences Education*, 13(3), 330-339. doi:10.1002/ase.1913


Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3181225>

Note: To cite this publication please use the final published version (if applicable).

Anatomy Dissection Course Improves the Initially Lower Levels of Visual-Spatial Abilities of Medical Undergraduates

Katerina Bogomolova, ^{1,2*} Beerend P. Hierck,^{2,3,4} Jos A. van der Hage,^{1,2} Steven E.R. Hovius^{5,6}

¹Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands

²Center for Innovation in Medical Education, Leiden University Medical Center, Leiden, The Netherlands

³Department of Anatomy and Embryology, Leiden University Medical Center, Leiden, The Netherlands

⁴Leiden Teachers' Academy, Leiden University, Leiden, The Netherlands

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

⁶Department of Plastic and Reconstructive Surgery, Radboud University Medical Center, Nijmegen, The Netherlands

Visual-spatial abilities are considered a successful predictor in anatomy learning. Previous research suggest that visual-spatial abilities can be trained, and the magnitude of improvement can be affected by initial levels of spatial skills. This case-control study aimed to evaluate (1) the impact of an extra-curricular anatomy dissection course on visual-spatial abilities of medical undergraduates and (2) the magnitude of improvement in students with initially lower levels of visual-spatial abilities, and (3) whether the choice for the course was related to visual-spatial abilities. Course participants ($n = 45$) and controls ($n = 65$) were first and second-year medical undergraduates who performed a Mental Rotations Test (MRT) before and 10 weeks after the course. At baseline, there was no significant difference in MRT scores between course participants and controls. At the end of the course, participants achieved a greater improvement than controls (first-year: $\Delta 6.0 \pm 4.1$ vs. $\Delta 4.9 \pm 3.2$; ANCOVA, $P = 0.019$, Cohen's $d = 0.41$; second-year: $\Delta 6.5 \pm 3.3$ vs. $\Delta 6.1 \pm 4.0$; $P = 0.03$, Cohen's $d = 0.11$). Individuals with initially lower scores on the MRT pretest showed the largest improvement ($\Delta 8.4 \pm 2.3$ vs. $\Delta 6.8 \pm 2.8$; $P = 0.011$, Cohen's $d = 0.61$). In summary, (1) an anatomy dissection course improved visual-spatial abilities of medical undergraduates; (2) a substantial improvement was observed in individuals with initially lower scores on the visual-spatial abilities test indicating a different trajectory of improvement; (3) students' preferences for attending extracurricular anatomy dissection course was not driven by visual-spatial abilities. *Anat Sci Educ* 13: 330–339. © 2019 American Association of Anatomists.

Key words: gross anatomy education; undergraduate education; visual-spatial abilities; cadaver dissection; spatial abilities; mental rotations test

INTRODUCTION

Anatomical education is constantly under pressure despite it being considered as one of the cornerstones of medical curricula. Teaching hours of anatomy have been decreasing over

time since the shift toward an integrated curriculum (Drake et al., 2002, 2009; Bergman et al., 2011). Additionally, ethical reasons, the high costs and limited availability of cadavers, and the increased time pressure on curricula 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, its educational value is under debate, dissection classes are found to be highly valuable by medical undergraduates, regardless of their sex, academic background, or citizenship (McLachlan et al., 2004; Azer and Eizenberg, 2007). In their opinion, dissections deepen their understanding of anatomical structures and their spatial relations, make learning interesting and are preferred over any other educational approach, especially in the

*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

Received 12 November 2018; Revised 5 July 2019; Accepted 8 July 2019.

Published online 24 July 2019 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/ase.1913

© 2019 American Association of Anatomists

first year of the medical program (Azer and Eizenberg, 2007). Today, medical undergraduates learn the anatomy mostly from two-dimensional (2D) representations of structures in anatomical atlases and textbooks and, consequently, experience difficulties to translate the acquired 2D knowledge into practice (McKeown et al., 2003; Prince et al., 2005; Spielman and Oliver, 2005; Bergman et al., 2008).

Visual-Spatial Abilities and Performance in Anatomy

How well acquired 2D anatomical knowledge is translated into practice depends largely on the visual-spatial abilities of students. In the medical anatomical context, it is defined as the ability that allows students to construct visual-spatial, e.g. three dimensional (3D), mental representations of 2D images and to mentally manipulate these representations (Gordon, 1986; Kozhevnikov and Hegarty, 2001). The first studies evaluating the association between visual-spatial abilities and anatomy learning have been performed by Rochford (1985) and Garg and colleagues (1999a,b; 2002). In these studies, visual-spatial abilities have significantly affected the learning process of spatial anatomy regardless of age, sex, right handedness, or computer use. Since then, even more research has been conducted to explore this association. The first comprehensive review of studies has been performed by Langlois and colleagues (2017). Their meta-analysis has revealed a predictive value of visual-spatial abilities when anatomy is assessed using spatial methods such as practical examination, 3D synthesis from 2D views, and drawing of views and cross-sections. As such, visual-spatial abilities are considered a successful predictor in anatomy learning and assessment (Yamine and Violate, 2015; Langlois et al., 2017). In health care professions, visual-spatial abilities are also a successful predictor in the acquisition of surgical technical skills, especially in the early stages of learning (Maan et al., 2012; Langlois et al., 2015a). For instance, Wanzel and colleagues (2003) have evaluated the correlation between visual-spatial abilities and surgical performance of dental students, surgical residents, and staff surgeons in performing a spatially complex surgical procedure. Visual-spatial abilities scores were correlated with surgical performance only within the group of dental students, suggesting that practice and surgical experience may supplant the influence of visual-spatial abilities over time. The effect of visual-spatial abilities on performance has also been demonstrated in mathematics (Hegarty and Kozhevnikov, 1999), veterinary education (Provo et al., 2002) and dental education (Hegarty et al., 2009).

Visual-Spatial Abilities as a Selection Tool

It is not surprising that visual-spatial abilities have been recommended to be used not only in the training, but also in the selection of surgical residents (Langlois et al., 2015a). A high motivation for the surgical specialty would apparently not be enough since it does not imply higher visual-spatial abilities among candidates. Langlois and colleagues (2015b) have evaluated a cohort of 210 medical graduates and did not find any relation between visual-spatial abilities and the choice of residency program. Nor did the choice for an elective course of applied anatomy depend on the visual-spatial abilities of medical graduates (Langlois et al., 2009). However, the relation between visual-spatial abilities and a high interest in anatomy,

in the very early stages of a medical career, has not yet been evaluated.

Malleability of Visual-Spatial Abilities

On the contrary, several studies have suggested that visual-spatial abilities can be trained through practice and experience. In a meta-analysis, Langlois and colleagues (2020) have found evidence for improvement of spatial abilities in anatomy education using instruction in anatomy and mental rotation training. For instance, in a single group study, Lufler and colleagues (2012) have reported an improvement of visual-spatial abilities of first-year medical undergraduates after participation in a gross anatomy course consisting of six dissection sessions. In a similar study with a control group of educational sciences students, visual-spatial abilities have increased after participation in the course consisting of lectures, self-study assignments including computer-assisted learning (CAL), collaborative learning, laboratory with prosected specimens, and body painting (Vorstenbosch et al., 2013). When an anatomy course was combined with a training of mental rotation skills unrelated to anatomy, an even higher increase in visual-spatial abilities scores has been observed (Hoyek et al., 2009). These were the only two studies to date that have included the practice effect on spatial abilities test scores in a control group resulting in a pooled treatment effect of 0.47 (95% CI [-0.03; 0.97]). The pooled treatment effect of single-group studies included in the meta-analysis was 0.49 (95% CI [0.17; 0.82], $n = 11$).

Furthermore, the improvement appears to be present on an expert level (Fernandez et al., 2011). It has been found that expert clinical anatomists were better in performing metric spatial tasks than novices, suggesting that visual-spatial abilities are trained by practice and education. In addition, the dose-dependent effect of practice and learning on visual-spatial abilities has been found in medical undergraduates after attending CAL courses of musculoskeletal and cardiovascular anatomy (Guimarães et al., 2019).

The malleability of visual-spatial abilities has been demonstrated in other disciplines as well, such as science, technology, engineering and mathematics (STEM), and veterinary medicine (Uttal et al., 2013; Gutierrez et al., 2017). In the meta-analysis of Uttal and colleagues (2013), visual-spatial abilities were classified as an intrinsic and dynamic spatial skill and were significantly affected by training with an overall effect size of 0.49 ($P < 0.01$).

Sex Differences and Initial Level of Performance

Sex differences in visual-spatial abilities have been repeatedly reported in the literature. At baseline, males have often achieved higher scores in visual-spatial abilities tests than females (Langlois et al., 2013; Uttal et al., 2013; Nguyen et al., 2014; Guimarães et al., 2019). This difference has been particularly observed in measures of mental rotation (Linn and Petersen, 1985; Peters et al., 2006). However, as has been demonstrated by several studies and meta-analyses, both males and females can achieve comparable magnitude of improvement after training (Baenninger and Newcombe, 1989; Uttal et al., 2013; Nguyen et al., 2014).

Another aspect worthy to mention is the initial level of performance of individuals in visual-spatial abilities training. A meta-analysis of 187 studies using a screening procedure

to identify initially low-performing students has reported significantly larger effect of training when compared to studies enrolling all participants regardless of initial performance levels (Uttal et al., 2013). These finding suggests that low-performing students can achieve a larger magnitude of improvement than high-performing students (Uttal et al., 2013). Additionally, students and residents with lower visual-spatial abilities in a surgical field have been able to achieve required levels of knowledge and skills through suitable teaching methods and guidance (Wanzel et al., 2003; Lufler et al., 2012; Cui et al., 2017; Roach et al., 2019). Therefore, it might be valuable to consider visual-spatial abilities as a tool to identify learners who will benefit most from extra practice and new learning environments (Yue, 2015) instead of an absolute selection criterium to guide selection of candidates for surgical training programs.

The Erasmus Medical Center Anatomy Research Project

The Erasmus MC Anatomy Research Project (EARP) is an extracurricular anatomy dissection course at the faculty of Medicine, Erasmus University Medical Center Rotterdam, The Netherlands. The EARP was set up in 2003 in response to reduced teaching volume of anatomy and a limited exposure to dissections. Since then, the course has become a unique and fully autonomous peer-to-peer educational model. The extracurricular course is organized annually during a period of 10 weeks. It takes place in the evening hours and does not interfere with the regular medical program. All medical undergraduates, from year one to year six of the undergraduate program, are invited to apply for one of the

four parallel programs, each covering a different anatomical region: Thorax (for the first-year students), Abdomen (for the second-year students), Head & Neck and Urogenital System (for the third-, fourth-, fifth-, and sixth-year students), and Extremities (for the third-, fourth-, fifth-, and sixth-year students). Due to a limited capacity, e.g. six available cadavers, a maximum of one hundred students are admitted annually, 24 students to Thorax, Abdomen and Extremities programs and 32 students to Head & Neck and Urogenital System program (Fig. 1). Student have to apply with a written assignment, e.g. about solving a clinical anatomy case. Selection of students is based on the highest scoring assignments and performed blindly by the EARP committee. After enrollment, students attend an instructional lecture and receive the EARP handbook with guidelines and detailed explanation of dissection of the assigned anatomical region including text and images. Subsequently, students start to work toward a complete dissection of the anatomical region on the assigned cadaver in a group of four students for eight weeks. Two students dissect the left part of the region, while the other two students dissect the right part of the region, which ensures equally active involvement of all students. Eventually, the same cadaver is used by four groups of two students each week, each group working on a different anatomical region on a different day of the week.

Dissection sessions are supervised by two tutors and four mentors who are senior medical undergraduates who previously participated in EARP. To ensure the quality of supervision and optimal knowledge of anatomy and dissection, all tutors and mentors attend a training program which also includes a dissection of the assigned anatomical regions. The EARP program includes 20-24 hours of dissections, 3-5 hours of interactive

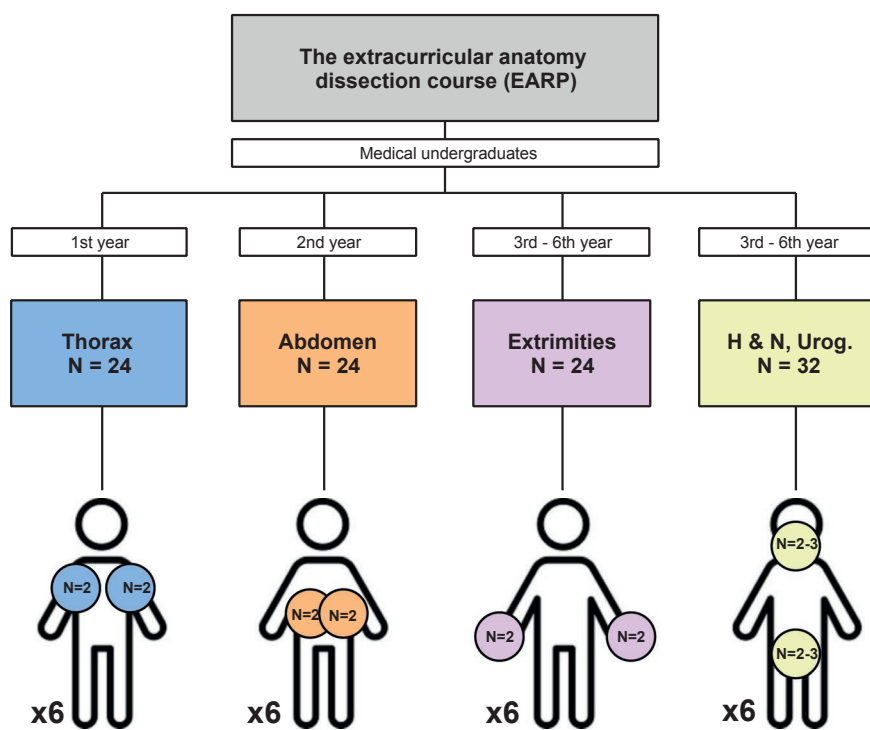


Figure 1.

An extra-curricular anatomy dissection course. Students attend eight dissection sessions of three hours each week. EARP, the Erasmus MC Anatomy Research Project; H & N, Urog., Head & Neck and Urogenital System.

lectures and demonstration sessions given by medical specialists and an hour of practical and written examinations. The latter is composed of questions assessing factual knowledge (e.g. naming a muscle's origin or insertion, innervation, and vascularization), spatial knowledge (e.g. the course of nerves and vessels in relation to other structures) and clinical decision making. The practical examination is composed of two parts: identification of as many structures as possible on a specimen during three minutes and naming of pin-pointed structures marked on a specimen.

Objectives and Aims

In the Netherlands, EARP has been established as a unique peer-to-peer educational setting in which spatial anatomy is learned hands-on during cadaveric dissections outside the regular medical program. This setting provides a unique opportunity to evaluate to what extent cadaveric dissection has its effect on visual-spatial abilities of medical undergraduates when compared to a control group consisting of non-participating medical undergraduates at the same stages of their curricula. In addition, it allows to evaluate a possible relation between having a high interest for anatomy and the visual-spatial abilities of students in the early phases of their medical careers.

Therefore, the aim of this study was to evaluate the impact of anatomy dissection course on visual-spatial abilities of medical undergraduates, and the magnitude of improvement in individuals with initially lower levels of visual-spatial abilities. Additionally, the present study aimed to evaluate whether the choice to apply for an extracurricular anatomy dissection course was related to the visual-spatial abilities of students. The authors hypothesized that individuals with higher levels of visual-spatial abilities are more likely to apply, that visual-spatial abilities will improve after an anatomy dissection course, and the improvement will be larger in individuals with initially lower levels of visual-spatial abilities.

MATERIALS AND METHODS

Study Design

A prospective case control-study was carried out at the Erasmus University Medical Center Rotterdam, the Netherlands. In general, a case-control study is efficient in evaluating associations between rare exposures and outcomes (Song and Chung, 2010). Since only 24 out of 400 students from each academic year participate in the EARP program, this study design was most suitable to answer the research questions. The study was approved by the course coordinator and the director of medical educational program and was considered exempt from formal assessment by the local ethical assessment committee (METC) of Erasmus University Medical Center Rotterdam (case number: CME-2019-0077).

Participants

Cases were defined as first-year and second-year medical undergraduates who were admitted to the EARP Thorax and Abdomen programs, respectively. Course participants, or cases, were identified through the attendance list of the programs. Controls were defined as first-year and second-year medical undergraduates who did not apply for the course and were matched for academic year and sex. Students who did

apply for the course, but were not selected, were excluded. For each course participant, a maximum of two controls were identified and approached during the regular lectures at the faculty with the request to participate. A 1 case : 2 control ratio was chosen since little is gained in terms of statistical power by including more than two controls for each case (Lewallen and Courtright, 1998).

Measurement of Spatial Abilities

Visual-spatial abilities were assessed by the Mental Rotations Test (MRT), previously validated by Vandenberg and Kuse (1978) which was based on rotated blocks of Shepard and Metzler (1971) and redrawn by Peters and colleagues (1995). This psychometric test is widely used in the assessment of visual-spatial abilities and has repeatedly showed a positive association with anatomy learning and assessment (Langlois et al., 2017). The test consists of a standard set of 24 items. Within each item, a three-dimensional (3D) figure is presented as a 2D drawing with four possible rotated versions of that figure. Subjects have to make a mental 3D representation and rotation of the figure in order to identify the two correct options. One point per item was awarded if both selected options were identified correctly. The maximum score on this test was 24 points.

A testing effect has been previously reported after repeated administration of the MRT (Hoyek et al., 2009; Martín-Gutiérrez et al., 2013; Vorstenbosch et al., 2013; Nguyen et al., 2014). In an attempt to minimize the testing effect, two versions of the MRT were used. The MRT, used as a pretest, included the original set of 24 items. In the MRT, used as a posttest, the same 24 items were rearranged in a different random order.

Procedures

Participation was voluntary, and an informed written consent was obtained by all participants before study. A short pre-questionnaire was used to gather information on age, sex, participation in EARP Thorax program in the first year (only applicable for second-year students) and prior or current participation in an academic program other than Medicine. A paper-and-pencil MRT pretest was administered to course participants prior to the start of their first dissecting session. The MRT posttest was performed after 10 weeks on the day of their examination. Controls simultaneously completed the MRT pretest and posttest in a lecture hall. All students were given 10 minutes to complete the test without a break.

Statistical Analysis

Descriptive statistics were used to summarize participants' baseline characteristics. Discrete variables were described as absolute frequencies (N) and percentages (%), and continuous variables as mean and standard deviation (SD). The differences in baseline characteristics were assessed with chi-squared test for differences in proportions and independent t-test for differences in means. The MRT scores were measured on a continuous scale and reported in terms of means and standard deviations. The differences in MRT pretest scores between course participants and controls were assessed with an independent t-test for normal distributions and Mann-Whitney test for non-parametric distributions. The differences in mean

improvement in MRT scores (Δ MRT) were assessed with a one-way ANCOVA. The mean improvement was included as a dependent variable, the EARP participation as a fixed factor and the absolute MRT pre-test score as a covariate. All analyses were adjusted for age, sex, participation in EARP Thorax program in the first year (only applicable for second-year students) and prior or current participation in an academic program other than Medicine. Additionally, the analysis was repeated for MRT-low (individuals who scored below the mean on the MRT pretest) and MRT-high (individuals who scored above the mean on the MRT pretest) groups separately with adjustment for academic year. Correlation between MRT pretest scores and mean improvement was assessed with Pearson correlation coefficient. The effect size (Cohen's *d*) of the differences in MRT improvement between groups was calculated using the mean scores and standard deviations of both 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

All course participants enrolled in the EARP Thorax and Abdomen programs participated in the study. For the 24 EARP Thorax participants a total of 44 controls were identified. For the 24 EARP Abdomen participants a total of 22 controls were identified. Four subjects were excluded from the analysis due to the following reasons: one participant selected only one correct option in the MRT pretest instead of two; two course participants did not complete the MRT posttest due to their absence on the EARP examination day; one control was a significant outlier and was removed from the analysis since a significant

outlier violates one of the required assumptions for performing a one-way ANCOVA and may reduce the validity of results.

Baseline Characteristics

No significant difference was found between course participants and controls in terms of age, sex and participation in an academic program other than Medicine (Table 1). The observed high ratio of females in both groups represents the average ratio of males and females in the current undergraduate medical curriculum in the Netherlands, which is approximately 30%:70%. The only significant difference was observed among second-year students in numbers of students who participated in the EARP Thorax program in the first year (nine students in the course participant group versus zero students in the control group, $P = 0.001$). The MRT pretest scores of these 9 students were not significantly different from the scores of the other 14 course participants (10.9 ± 4.3 vs. 12.3 ± 6.4 , $P = 0.272$).

Improvement in Mental Rotations Test Scores

As shown in Figure 2, no significant difference in MRT pretest scores was found between the course participants and controls (*first-year*: 14.6 ± 5.5 vs. 13.8 ± 5.9 ; $P = 0.411$; *second-year*: 11.8 ± 5.1 vs. 11.5 ± 5.2 ; $P = 0.856$). After 10 weeks, the MRT scores were significantly improved in both the groups. However, the mean improvement (Δ MRT) among course participants was significantly higher than the controls (*first-year*: $\Delta 6.0 \pm 4.1$ vs. $\Delta 4.9 \pm 3.2$; $F_{(1,56)} = 5.8$, $P = 0.019$, Cohen's $d = 0.31$; *second-year*: $\Delta 6.5 \pm 3.3$ vs. $\Delta 6.1 \pm 4.0$; $F_{(1,36)} = 2.7$, $P = 0.03$, Cohen's $d = 0.11$) (Fig. 2). Higher MRT pretest scores were associated with less improvement in both academic years

Table 1.

Baseline Characteristics of Study Participants

Characteristics	Course participants	Controls	P-value
First year	n = 22	n = 43	
Age, mean (\pm SD)	19.3 (\pm 1.4)	18.7 (\pm 1.0)	0.587
Sex			
Male, n (%)	8 (27.3)	9 (20.9)	0.569
Female, n (%)	16 (72.7)	34 (79.1)	
Participation in an academic program other than Medicine, n (%)	4 (18.2)	2 (4.7)	0.084
Second year	n = 23	n = 22	
Age, mean (\pm SD)	19.4 (\pm 4.4)	19.8 (\pm 1.1)	0.546
Sex			
Male, n (%)	6 (26.1)	5 (22.7)	0.799
Female, n (%)	17 (73.9)	17 (77.3)	
Participation in an academic program other than Medicine, n (%)	1 (4.3)	4 (18.2)	0.187
Participated in EARP Thorax program in the first year, n (%)	9 (39.1)	0 (0)	0.001

n, number of students; \pm SD, \pm Standard Deviation; EARP, Erasmus MC Anatomy Research Project.

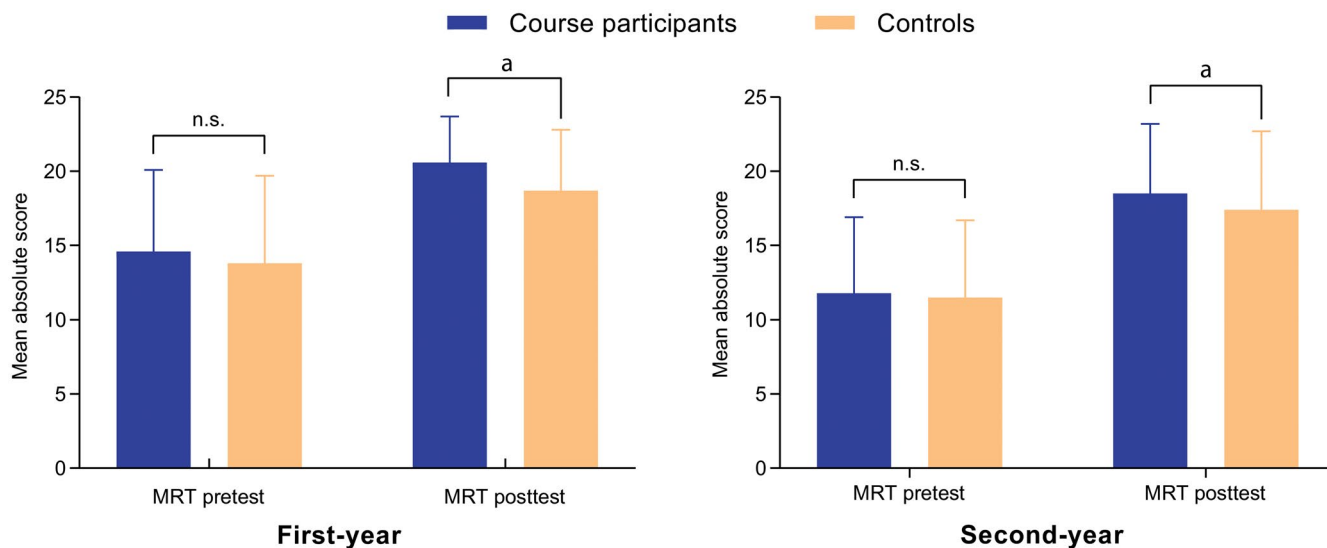


Figure 2.

Differences in performance on Mental Rotations Test (MRT) pretest and MRT posttest between course participants [first year: the Erasmus MC Anatomy Research Project (EARP) Thorax program, second year: EARP Abdomen program] and controls. Performances are reported in mean scores. Error bars represent standard deviation; ^a $P < 0.05$; n.s., not significant; MRT, Mental Rotations Test.

(first-year: $\beta = -0.9$; 95% CI [-1.3; -0.3], $P = 0.0001$; second year: $\beta = -0.3$; 95% CI [-0.52; -0.14], $P = 0.001$). Additionally, among second-year students, previous participation in EARP was negatively associated with the mean improvement in MRT scores ($\beta = -3.9$; 95% CI [-1.16; -6.68], $P = 0.07$). Sex, age, and participation in an academic program other than Medicine were not significantly associated with the improvement.

Improvement in Students with Initially Lower Mental Rotations Test Scores

As shown in Figure 3, when the analysis was repeated for individuals who scored below and above average on the MRT pretest (e.g. MRT-low and MRT-high groups) separately, the improvement in MRT scores was only present in the MRT-low group with a much larger effect size (*MRT-low group*: $\Delta 8.4 \pm 2.3$ vs. $\Delta 6.8 \pm 2.8$; $F_{(1,50)} = 6.916$, $P = 0.011$, Cohen's $d = 0.61$; *MRT-high group*: $\Delta 3.8 \pm 3.3$ vs. $\Delta 3.6 \pm 2.7$; $F_{(1,45)} = 1.253$, $P = 0.269$, Cohen's $d = 0.06$).

Additionally, as shown in Figure 4, the negative association between MRT pretest scores and mean improvement in MRT scores was no longer present. Instead, course participants in the MRT-low group showed a positive correlation between MRT pretest scores and mean improvement ($r = 0.350$, $P = 0.093$). In the MRT-high group, however, around 55% ($R^2 = 0.55$) of the total variation in MRT posttest scores could be explained by the MRT pretest scores. There was a moderate negative correlation between mean improvement and MRT pretest scores in course participants ($r = -0.68$, $P = 0.001$) and controls ($r = -0.76$, $P = 0.001$).

Sex Differences

Males significantly outperformed females on the MRT pretest (15.2 ± 5.8 vs. 12.5 ± 5.3 , $P = 0.034$) and on the MRT posttest

(20.2 ± 4.1 vs. 18.4 ± 4.2 ; $P = 0.038$). However, there was no significant difference in the mean improvement (Δ MRT) between males and females ($\Delta 5.04 \pm 4.0$ vs. $\Delta 5.9 \pm 3.2$; $F_{(1,100)} = 0.371$, $P = 0.962$). Additionally, the percentage of females in the MRT-low group did not differ significantly from the percentage in the MRT-high group (84.2% vs. 70.0%, $\chi^2 = 3.091$, $P = 0.079$).

DISCUSSION

This case-control study was performed to evaluate the impact of an extra-curricular anatomy dissection course on the visual-spatial abilities of medical undergraduates and to evaluate whether the choice for this course was related to the initial level of their visual-spatial abilities. Furthermore, a control group composed of medical undergraduates was included which enhances the internal and external validity of the results. The study resulted in the following findings and observations.

First, the results of this experimental study showed a significant improvement of visual-spatial abilities, as measured by the Mental Rotations Test, after completing eight sessions of cadaveric dissections. The observed effect sizes (first-year: $d = 0.31$; second-year: $d = 0.11$) indicate that a repeated practice of dissection had a small to medium effect on the visual-spatial abilities of students. This effect was much smaller than the one observed by Lufner and colleagues (2012) ($d = 1.02$) after a dissection course when no control group was included (Langlois et al., 2020). The difference in effect sizes can be attributed to the testing effect. This effect occurs after repeated administration of the MRT which provides students the chance to train their spatial skills by doing the test (Peters et al., 1995; Geiser et al., 2006; Hegarty et al., 2009). This practice effect in the current study was reflected by a significant improvement of the MRT scores among controls. A similar effect has also been observed by Vorstemsbosch and colleagues by including a control group composed of students of educational sciences

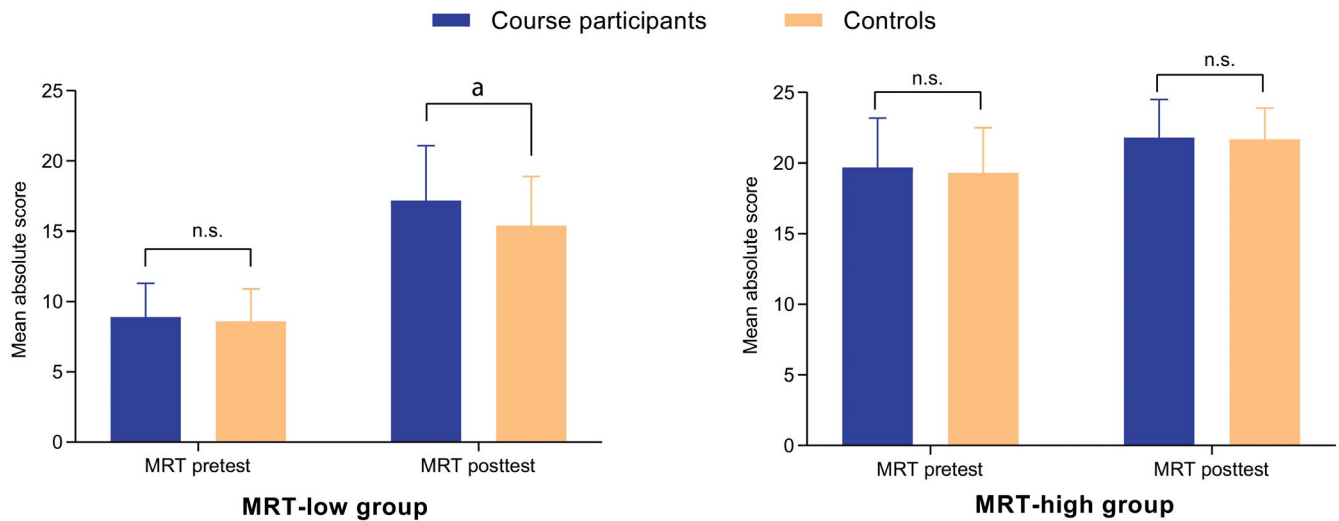


Figure 3.

Differences in performance on Mental Rotations Test (MRT) pretest and MRT posttest between course participants and controls in the MRT-low and MRT-high groups of first- and second-year medical undergraduates. Performances are reported in mean scores. Error bars represent standard deviation; ^a $P < 0.05$; n.s., not significant; MRT, Mental Rotations Test.

(Vorstenbosch et al., 2013). A control group is, therefore, essential for the assessment of the course-related improvement. Additionally, a control group composed of the identical source population, e.g. medical undergraduates from the same academic years, eliminate possible differences in baseline characteristics, such as high school profiles, intellectual interest and hobbies, that can influence visual-spatial abilities.

The improvement in visual-spatial abilities scores may be attributed to active involvement in dissection of a 3D cadaver accompanied with studying 2D representations from the EARP handbook. Additionally, students are constantly challenged by mental visualization of anatomical structures and understanding of their spatial relations in order to perform the dissection in the best and most efficient way. Further research is needed to determine which components of cadaveric dissection contribute most to the improvement in visual-spatial abilities, and to what extent this effect will remain present. In the current study, nine second-year course participants, participated in the EARP Thorax program in the previous year. They did not perform better on the MRT pretest test than the 14 course participants, who participated in the EARP course for the first time. This may suggest that the acquired level of visual-spatial abilities might not be long lasting. However, this sample was too small to draw that conclusion.

Second, when the results were analyzed for individuals with initially lower MRT pretest scores only, a much larger effect size ($d = 0.61$) was observed. Individuals with initially higher MRT pretest scores did not show any improvement. Instead, the MRT pretest scores were negatively associated with the mean improvement. These findings may reflect an aptitude-treatment effect of visual-spatial abilities, i.e. low-performing individuals are having a different trajectory of improvement than high-performing individuals (Cook, 2005). That visual-spatial abilities may cause an aptitude-treatment interaction has been illustrated earlier by Cui and colleagues (2017). After learning with monoscopic 3D images, students with lower visual-spatial abilities scores performed significantly worse than students with higher visual-spatial abilities scores. While after

learning with a stereoscopic 3D model, these students performed significantly better and equally well as students with higher visual-spatial abilities scores. Similar effects have been reported by Garg and colleagues (1999b; 2002) where students with lower visual-spatial abilities had significant disadvantages by learning anatomy with multiple view presentations, while students with higher visual-spatial abilities performed better with these types of presentations (Garg et al., 1999b, 2002). The observed phenomenon in this study, however, may also be attributed to the ceiling effect in the MRT test. This effect is addressed further in the limitation section.

Third, the choice for an extracurricular anatomy dissection course, in this study, did not imply higher levels of visual-spatial abilities of medical undergraduates. These findings support previous research on spatial abilities and the choice for an elective course of applied anatomy (Langlois et al., 2009) or personal preference for a surgical specialty (Langlois et al., 2015b). In both situations, personal preferences and choices among postgraduates were not reported to be associated with the individual visual-spatial abilities. It is interesting to note that the choice for medical careers in the first place may imply higher visual-spatial abilities among medical undergraduates. When compared to students of educational sciences, medical undergraduates had higher mean visual-spatial abilities (Vorstenbosch et al., 2013). Similar differences were observed between dental and psychology students (Hegarty et al., 2009).

Lastly, the observed sex differences in this study were in line with the previous research. Despite having initially lower scores on the MRT test, females were still able to achieve similar magnitudes of improvement as males after training (Baenninger and Newcombe, 1989; Langlois et al., 2013; Uttal et al., 2013; Nguyen et al., 2014; Guimarães et al., 2019). Additionally, in this study, the percentage of females in the MRT-low group was not significantly lower than in the MRT-high group, as could be expected. These findings suggest that the individual approach is preferable, since a particular male may have lower visual-spatial abilities than a particular female.

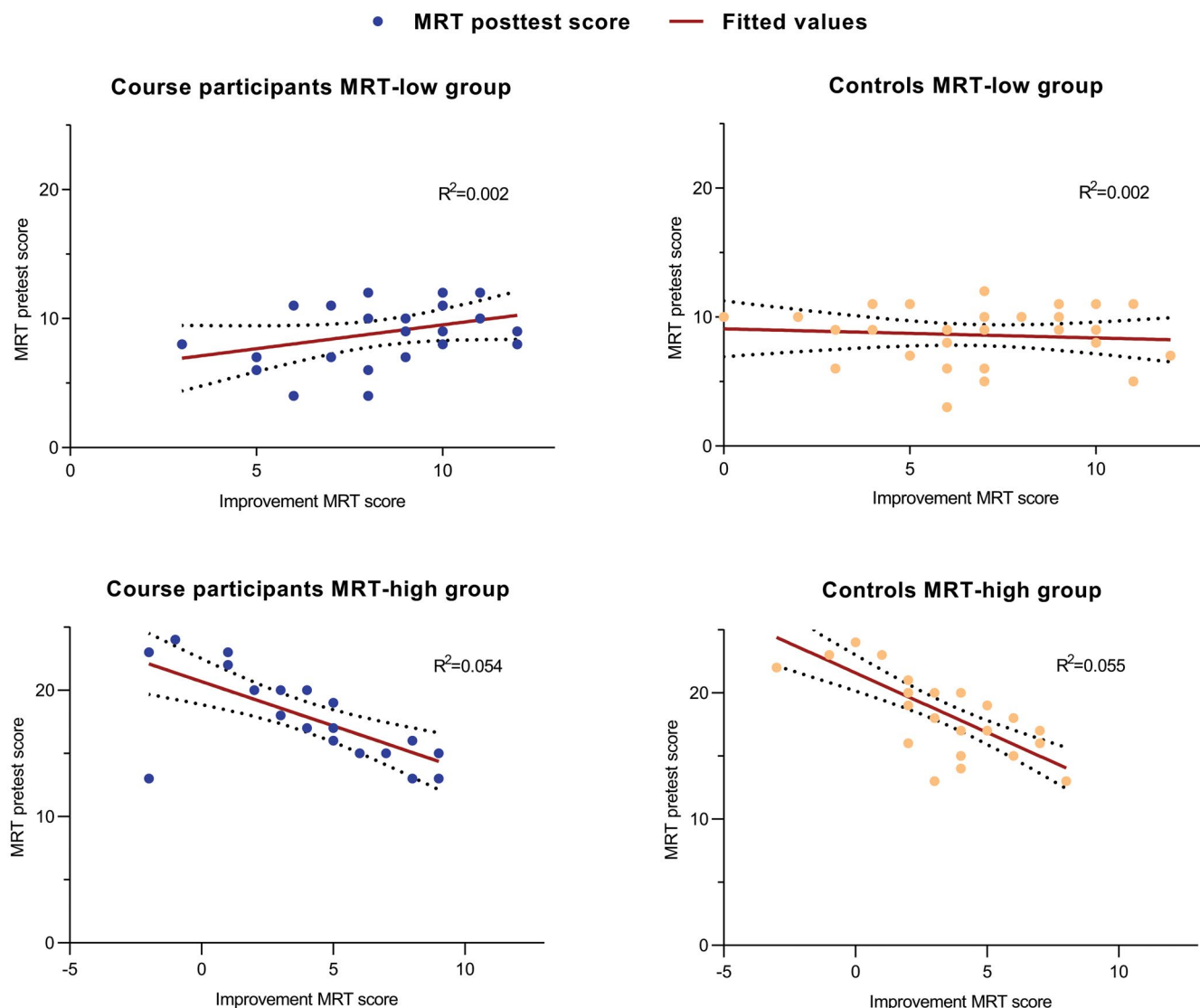


Figure 4.

Relationship between Mental Rotations Test (MRT) pretest scores and mean improvement in the MRT scores. A regression analysis graph illustrating the relationship between initially low and high levels of visual-spatial abilities and mean improvement among course participants and controls. MRT, Mental Rotations Test.

Future Directions

The findings of this study underline the importance of anatomical education in the light of visual-spatial abilities training. The positive effect of anatomical education on visual-spatial abilities, which in turn facilitates learning and retention of anatomical knowledge, indicates that these two can reinforce one another. Additionally, Roach and colleagues have demonstrated that an early guidance and instruction can improve low-performing students' strategies for spatial problem-solving (Roach et al., 2017a,b; 2019). This can be of a great importance for low-performing individuals and have implications for individualized approaches in the current curricula.

The role of augmented and virtual reality in anatomical education is promising and is currently addressed in ever more research. In the fields of engineering and technology, research has shown that training in augmented and virtual reality can

improve various components of spatial abilities, such as visualization, rotation, and orientation (Hegarty et al., 2009; Hoyek et al., 2009; Martín-Gutierrez et al., 2013; Roca-González et al., 2017). Stereoscopic 3D visualization technologies may, therefore, serve as valuable additional tools to include spatial reasoning training in an anatomical context next to traditional ways of learning (Wainman et al., 2018).

Limitations of the Study

A case control study is relatively quick and efficient in evaluating associations between rare exposures and outcomes (Song and Chung, 2010). Since only 24 out of 400 students participate in the extracurricular anatomy dissection course in each academic year, the numbers of participants were restricted. Since this study design requires comparatively few subjects, it

allowed to omit recruitment of the entire first and second-year cohorts. However, a desirable 1:2 case:control ratio among second-year students was not achieved. To underline the validity of 1:1 case:control ratio in the main analysis, a post hoc analysis among first-year students was performed. After random elimination of half of the controls and repeated analysis, a significant difference between course participants and controls remained (20.6 vs. 18.5; ANCOVA, $F = 4.8$, $P = 0.034$; Cohen's $d = 0.09$). Consequently, the recruited amount of controls among second-year students was justifiable for the main analysis.

Certainly, a case-control study is susceptible to particular types of biases. Since no randomization was possible in this setting, a selection bias should be taken into account despite of the recruitment of the controls from the identical source population. Controls were recruited in the lecture hall and only part of them was willing to participate in the study. They could have been less motivated to do their best on the MRT test than the course participants who were usually highly motivated and were more willing to perform best on such a test. This could have partially accounted for the less improvement in the MRT scores among controls. Other possible confounders, which were not included, were gaming experience and performance on anatomy in the current curriculum. Both have been associated with a better performance in visual-spatial abilities tests before (Terlecki 2008; Uttal et al., 2013).

The MRT pretest scores were negatively associated with the mean improvement, especially in the MRT-high group. This association may reflect an aptitude-treatment interaction, but a ceiling effect cannot be ruled out. Ceiling effect occurs when more than 15% of the participants reach the highest possible scores of a test (Lim et al., 2015). In this study, 13% of the participants reached the highest possible score of 24 points. Therefore, a ceiling effect was not likely but cannot be ruled out completely. To avoid a possible ceiling effect in the future, a more difficult set of items in the MRT could be used allowing high-performing students achieve a much greater improvement. The association could also be attributed the statistical feature "regression to the mean", i.e. since high-performing students structurally score higher on the pretest, they are more likely to score lower on a repeated test.

CONCLUSIONS

This study showed that the visual-spatial abilities scores of medical undergraduates improved after anatomy dissection. Additionally, a substantial improvement was observed in individuals with initially lower scores on the visual-spatial abilities test. Although a ceiling effect cannot be completely ruled out, this can be indicative of a different trajectory of improvement between individuals in this particular study. This possible aptitude-treatment effect will need to be evaluated in further research and an individualized approach in current curricula could be considered. Finally, the students' preferences for attending the extracurricular anatomy dissection course were not driven by visual-spatial abilities.

ACKNOWLEDGMENTS

The authors sincerely thank Prof. Dr. Gert-Jan Kleinrensink for the collaboration and making this research project possible. 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.

BEEREND P. HIERCK, Ph.D., is an associate professor of anatomy in the Department of Anatomy and Embryology and 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.

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.

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, Nijmegen, The Netherlands. He teaches anatomy of upper and lower limbs to medical students and surgical residents.

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.
- Baenninger M, Newcombe N. 1989. The role of experience in spatial test performance: A meta-analysis. *Sex Roles* 20:327–344.
- 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, van der Vleuten CP, Scherpbier AJ. 2011. Why don't they know enough about anatomy? A narrative review. *Med Teach* 33:403–409.
- Bergman EM, de Bruin AB, Herrler A, Verheijen 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.
- Cohen J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Ed. Hillsdale, NJ: Lawrence Earlbaum 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.
- 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, Lowrie DJ, Prewitt CM. 2002. Survey of gross anatomy, microscopic anatomy, neuroscience, and embryology courses in medical school curricula in the United States. *Anat Rec* 269:118–122.
- 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.
- Fernandez R, Dror IE, Smith C. 2011. Spatial abilities of expert clinical anatomists: Comparison of abilities between novices, intermediates, and experts in anatomy. *Anat Sci Educ* 4:1–8.
- 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.
- Geiser C, Lehmann W, Eid M. 2006. Separating "rotators" from "nonrotators" in the mental rotations test: A multigroup latent class analysis. *Multivariate Behav Res* 41:261–93.

- Gordon HW. 1986. The cognitive laterality battery: Tests of specialized cognitive function. *Int J Neurosci* 29:223–244.
- 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.
- Gutierrez JC, Chigerwe M, Ilkiw JE, Youngblood P, Holladay SD, Srivastava S. 2017. Spatial and visual reasoning: Do these abilities improve in first-year veterinary medical students exposed to an integrated curriculum? *J Vet Med Educ* 44:669–675.
- Hegarty M, Kozhevnikov M. 1999. Types of visual-spatial representation and mathematical problem solving. *J Educ Psychol* 91:684–689.
- Hegarty M, Keehner M, Khooshabeh P, Montello DR. 2009. How spatial abilities enhance, and are enhanced by, dental education. *Learn Indiv Differ* 19:61–70.
- Hoyek N, Collet C, Rastello O, Fargier P, Thiriet P, Guillot A. 2009. Enhancement of mental rotation abilities and its effect on anatomy learning. *Teach Learn Med* 21:201–206.
- Kozhevnikov M, Hegarty M. 2001. A dissociation between object manipulation spatial ability and spatial orientation ability. *Mem Cognit* 29:745–756.
- Langlois J, Bellemare C, Toulouse J, Wells GA. 2015a. Spatial abilities and technical skills performance in health care: A systematic review. *Med Educ* 49:1065–1085.
- Langlois J, Bellemare C, Toulouse J, Wells GA. 2017. Spatial abilities and anatomy knowledge assessment: A systematic review. *Anat Sci Educ* 10:235–241.
- Langlois J, Bellemare J, Toulouse J, Wells GA. 2020. Spatial abilities training in anatomy education: A systematic review. *Anat Sci Educ* 13:71–79.
- Langlois J, Wells GA, Lecourtois M, Bergeron G, Yetisir E, Martin M. 2009. Spatial abilities in an elective course of applied anatomy after a problem-based learning curriculum. *Anat Sci Educ* 2:107–112.
- Langlois J, Wells GA, Lecourtois M, Bergeron G, Yetisir E, Martin M. 2013. Sex differences in spatial abilities of medical graduates entering residency programs. *Anat Sci Educ* 6:368–375.
- Langlois J, Wells GA, Lecourtois M, Bergeron G, Yetisir E, Martin M. 2015b. Spatial abilities of medical graduates and choice of residency programs. *Anat Sci Educ* 8:111–119.
- Lewallen S, Courtright P. 1998. *Epidemiology in practice: Case-control studies*. *Community Eye Health* 11:57–58.
- Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. 2015. Floor and ceiling effects in the OHS: An analysis of the NHS PROMs data set. *BMJ Open* 5:e007765.
- Linn MC, Petersen AC. 1985. Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Dev* 56:1479–1498.
- Lufler RS, Zumwalt AC, Romney CA, Hoagland TM. 2012. Effect of visual-spatial ability on medical students' performance in a gross anatomy course. *Anat Sci Educ* 5:3–9.
- Maan ZN, Maan IN, Darzi AW, Aggarwal R. 2012. Systematic review of predictors of surgical performance. *Br J Surg* 99:1610–1621.
- Martín-Gutiérrez J, Navarro-Trujillo RE, Acosta-Gonzalez MM. 2013. Augmented reality application assistant for spatial ability training. HMD vs computer screen use study. *Procedia Social Behav Sci* 93:49–53.
- 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.
- McLachlan JC, Bligh J, Bradley P, Searle J. 2004. Teaching anatomy without cadavers. *Med Educ* 38:418–424.
- Nguyen N, Mulla A, Nelson AJ, Wilson TD. 2014. Visuospatial anatomy comprehension: The role of spatial visualization and problem-solving strategies. *Anat Sci Educ* 7:280–288.
- 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.
- Peters M, Lehmann W, Takahira S, Takeuchi Y, Jordan K. 2006. Mental rotation test performance in four cross-cultural sample (n = 3367): Overall sex differences and the role of academic program in performance. *Cortex* 42:1005–1014.
- 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.
- Provo J, Lamar C, Newby T. 2002. Using a cross-section to train veterinary students to visualize anatomical structures in three dimensions. *J Res Sci Teach* 39:10–34.
- Pryde FR, Black SM. 2005. Anatomy in Scotland: 20 years of change. *Scott Med J* 50:96–98.
- 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.
- Roca-González C, Martín-Gutiérrez J, García-Domínguez M, del Carmen Mato Carrodegas M. 2017. Virtual technologies to develop visual-spatial ability in engineering students. *Eurasia J Math Sci Tech Educ* 13:441–468.
- Rochford K. 1985. Spatial learning disabilities and underachievement among university anatomy students. *Med Educ* 19:13–26.
- Shepard RN, Metzler J. 1971. Mental rotation of three-dimensional objects. *Science* 171:701–703.
- Song JW, Chung KC. 2010. Observational studies: Cohort and case-control studies. *Plast Reconstr Surg* 126:2234–2242.
- Spielmann PM, Oliver CW. 2005. The carpal bones: A basic test of medical students and junior doctors' knowledge of anatomy. *Surgeon* 3:257–259.
- Terlecki MS, Newcombe NS, Little M. 2008. Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Appl Cognit Psychol* 22:996–1013.
- Uttal DH, Meadow NG, Tipton E, Hand LL, Alden AR, Warren C, Newcombe NS. 2013. The malleability of spatial skills: A meta-analysis of training studies. *Psychol Bull* 139:352–402.
- Vandenberg SG, Kuse AR. 1978. Mental rotations, a group test of three-dimensional spatial visualization. *Percept Mot Skills* 47:599–604.
- Vorstenbosch MA, Klaassen TP, Donders AR, Kooloos JG, Bolhuis SM, Laan RF. 2013. Learning anatomy enhances spatial ability. *Anat Sci Educ* 6:257–262.
- 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.
- Wanzel KR, Hamstra SJ, Caminiti MF, Anastakis DJ, Grober ED, Reznick RK. 2003. Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. *Surgery* 134:750–757.
- Waterston SW, Stewart IJ. 2005. Survey of clinicians' attitudes to the anatomical teaching and knowledge of medical students. *Clin Anat* 18:380–384.
- Yamine 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.
- Yue C. 2015. Predicting and influencing training success: Spatial abilities and instructional design. *Med Educ* 49:1054–1055.