From bones to bytes: Do manipulable 3D models have added value in osteology education compared to static images?

Vicky Vandenbossche[1], Martin Valcke[2], Wouter Willaert[1], Emmanuel Audenaert[1,3,4,5]

Affiliations

[1] Department of Human Structure and Repair, Ghent University, Corneel Heymanslaan 10, 9000 Ghent, Belgium

[2] Department of Educational Studies, Ghent University, Ghent, Belgium

[3] Department of Orthopedic Surgery and Traumatology, Ghent University Hospital, Corneel Heymanslaan 10, 9000 Ghent, Belgium

[4] Department of Trauma and Orthopedics, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge, CB2 0QQ, UK

[5] Department of Electromechanics, Op3Mech research group, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium

Corresponding Author: Vicky Vandenbossche, Ghent University, Corneel Heymanslaan 10, 9000 Ghent, Belgium. E-mail: vicky.vandenbossche@ugent.be

ABSTRACT

Background: Over the past few years, anatomy education has been revolutionized through digital media, resulting in innovative computer-based 3D models to supplement, or even replace traditional learning materials. However, the added value of these models in terms of learning performance remains unclear. Multiple mechanisms may contribute to the inconclusive findings. This study focusses on the impact of active manipulation on learning performance and the influence that posttest design features may have on the outcome measurement.

Methods: Participants were randomly assigned to one of two research conditions: studying on the base of a computer-based manipulable pelvic bone model versus online static images of the same model. Pretests focused on students' baseline anatomy knowledge and spatial ability. Three knowledge posttests were administered: a test based on a physical pelvic bone model, and two computer-based tests based on static images and a manipulable model. Mental effort was measured with the Paas mental effort rating scale.

Results: In the static-images-based posttest, significantly higher knowledge scores were attained by participants studying in the static images research condition ($p = 0.043$). No other significant knowledge related differences could be observed. In the manipulable-model-based posttest, spatial ability rather than the research condition seemed to have an influential role on the outcome scores ($r = 0.18$, $p = 0.049$). Mental effort scores reflected no difference between both research conditions.

Conclusion: The research results are counter-intuitive; especially because no significant differences were found in the physical-model-based posttest in students who studied with the manipulable model. Explaining the results builds on differences in anatomical models

requiring less or more active manipulation to process spatial information. The pelvic bone manipulable model, and by extension osteology models, might be insufficiently complex to provide added value compared to static images. Moreover, the posttest modality should be chosen with care since spatial ability rather than anatomy knowledge may be measured.

INTRODUCTION

Background

Anatomy learning based on drawings, cadaver dissections and prosections has long been the sole way for students to appreciate three-dimensional (3D) relationships in the human body. More recently, anatomy education has been revolutionized through digital media, resulting in innovative computer-based 3D models (3DCBMs) to supplement, or even replace traditional learning materials. These 3DCMs may provide teachers with opportunities to tackle the current challenges (e.g. time constraints, lack of trained faculty, rising costs, etc.) in anatomy education.[1] Moreover, digital models became vital during the COVID-19 outbreak, accelerating their development which resulted in major advances in realism, scalability and user satisfaction.[2–4]

The variety of 3DCBMs used in anatomy education is large, ranging from desktop-based to fully immersive models.[5] However, if these models are going to be widely adopted, it is important that they are designed appropriately in order to mitigate abundant cognitive load. Therefore, Birbara and colleagues proposed instructional design guidelines that are grounded in the Cognitive Load Theory (CLT).[6] The delivery modality is one factor that needs to be considered during the design of the 3DCBM.[6] This includes interactivity with the models, the level of immersion, stereopsis, and motion.[6] To make inferences about the value of these factors, researchers have tried to investigate these factors individually.[7] A few studies have focused on the influence of interactivity.[8–11] Although it would seem that presenting virtual objects – i.e., manipulable 360° in all three planes - would have clear advantages over static atlas-type images, research findings do not support this assumption. In a study examining the learning of wrist bone anatomy, multiple-view computer models did not have an advantage over flat, key view (top, bottom, front, back and sides) representations; occasionally even disadvantages have been reported in learners with poor spatial ability.[8] Follow-up studies support these findings, but suggest and underpin the potential of offering additional views.[9–11] Studies focusing on a fully manipulable model did not report conclusive results either.[12–14]

The empirical evidence contradicts hypotheses based on the embodied cognition theory. Proponents of embodied cognition postulate that bodily interactions with the environment contribute to cognition.[15] Shepard & Metzler's seminal research showed that people mentally manipulate objects similarly to the way they would with actual objects, and that the time it takes to rotate the image increases linearly with the degree of rotation.[15,16] Additionally, Kosslyn and colleagues have shown that earlier manual manipulation of an object makes motor systems activity stronger when performing a mental rotation of that object due to a stronger embodied representation.[17] Furthermore, embodied representations could play a central role in anatomy instruction. Anatomy research has demonstrated that perception of real life anatomical models primes an embodied approach. One is inclined to relate observed anatomical models to one's own body coordination systems.[18] Therefore, manipulating anatomical models in a virtual environment is believed to result in the development of stronger "embodied", multi-modal mental representations.

The conflicting results might be compounded by a number of factors.

Firstly, there is much heterogeneity in the delivery modality of the models used in previous studies, e.g., multiple view models with rotation at 10° intervals [8] versus models manipulable 360° in all three planes.[12] This makes it hard to make inferences about the educational value of these models since these differences influence the level of manipulation, and consequently

- according to embodied cognition principles and the Cognitive Load Theory - the level of impact on cognitive representations.[6]

Secondly, many of the studies cited above have confounded research designs. To help unravel these confounded research designs, David A. Cook issued some guidelines.[19] To begin, media-comparison studies – e.g., comparing computer-based learning with noncomputer instruction (e.g. paper-based, face-to-face) – are confounded. In addition, Cook defined three instructional design levels: 1) the configuration (CD-ROM-based, Web-page-based, etc.), 2) the instructional method (self-assessment questions, clinical cases, etc.), and 3) the presentation (e.g. font, interaction, stereopsis, motion, etc.). To create generalizable results, Cook recommended developing research conditions within only one of these levels.[19]

Finally, assessment approaches are important to consider, since a lack of constructive alignment could be present.[20] With emerging 3D learning initiatives, misalignment of learning with assessment is growing. While there is a significant shift in teaching methods from 2D to 3D methods, the assessment methods are shifting the opposite way. Today, medical students are more frequently assessed using written tests, such as multiple choice questions (MCQ), extended matching questions (EMQ), and single best answer questions (SBA), which are 2D in nature.[21] Therefore, it is important to take into account the influence of the assessment modality on student performance in order that no bias enters modality preference.

Study aim & hypotheses

To address the above concerns, a research study was designed to explore the mechanism of active manipulation of 3DCBMs, as well as the influence of posttest design features on the outcome measurement. Special emphasis was put on the recommendations of Cook to execute research within one medium and within one level of instructional design.[20]

To assess this, learning outcomes of students learning with a computer-based manipulable model (MM) – *defined for our purposes as a virtual object manipulable 360° in all three planes on a computer screen in the same manner as a real object* – were compared with learning outcomes of students learning with online static images (SI) of the same model. Three different testing modalities were used: 1) a test building on a physical model, 2) a test building on an online MM, and 3) a test building on online SI. Moreover, the mental effort was measured.

Following hypotheses were put forward:

1) Based on embodied cognition principles and the CLT,[6,18] learning with an online MM will result in better learning outcomes on the test based on the physical model compared to learning with online SI.
2) Based on the CLT that postulates physical demands decrease when dealing with familiar material,[6] it is hypothesized that students who learn with a) the MM will perform better on a test building on the MM, and students learning on the base of b) the SI will perform better on a test building on SI.
3) Based on the CLT and instructional design principles,[6] the self-reported mental effort will be lower in the MM condition.

Methods

Research participants

All first-year medical students of Ghent University, Belgium (n = 453) were invited to participate on a voluntary base. The study took place at the start of the second semester. At this stage, students were not expected to have any previous knowledge of musculoskeletal anatomy. The study protocol was approved by the local Ethical Committee (2019/1935), and after recruitment informed consent was obtained. Participants were randomly (simple randomization) assigned to one of two research conditions: SI or MM.

Study design

As depicted in Figure 1, a research study was designed to explore the mechanism of learning with a manipulable model versus static images, as well as the influence of assessment modality on the outcome measurement.

Intervention

The learning phase was based on a pelvic bone model. Due to its relative spatial complexity, this bone was considered to be representative for other bones in the human body. The MM was created using a high resolution (0.05 mm) structured-light 3D surface scanner, the Artec® Space Spider (Artec 3D, Luxembourg, Luxembourg). Following the creation, the 3D model was uploaded and 26 anatomical structures were labeled within Sketchfab 3D rendering platform (Sketchfab, New York, NY, USA). Sketchfab is a web-based viewing, creating and publishing tool for 3D models. It offers a universal 3D/virtual reality (VR) viewer that can be accessed on any browser. The published pelvic bone model was then embedded into the university-specific learning platform (D2L/Brightspace, Kitchener, ON, Canada) and could be displayed on any laptop. The model was manipulable, allowing panning, rotation, and zooming via a mouse or touchpad.

The SI were based on print screens of the MM. Print screens were taken from the anterior, medial and dorsal side of the pelvic bone. The model was labelled similar to the MM. The SI were also embedded into the university-specific learning platform. Participants were able to consult the different views at their discretion.

During the learning phase, students were provided 15 minutes to study the 26 anatomical structures with their respective models based on a series of labels affixed to the appropriate regions. The time was chosen based on previous studies[12,22] to ensure that there was a range of observed scores, with neither ceiling nor basement effects.

Data collection

At the start of the experiment (T0), all participants completed a set of questionnaires, consisting of an anatomy knowledge test, a spatial ability test *(cfr. § Covariates),* and standard demographic questions. The anatomy test consisted of 10 nominal questions with a randomized word bank to assist in spelling and recall. This test helped establishing a baseline value about students' prior knowledge of the pelvic bone and to ascertain whether both research groups were comparable.

After the intervention (T1), i.e., the 15 minutes of learning, all students were asked to complete a second series of research instruments. First, a mental effort scale was administered (*cfr. § Covariates)*. Secondly, a 10-item anatomy knowledge test (4 orientation questions and 6 nominal questions) on a physical bone was administered. This test was included to study the influence of learning modality and to ensure that there was no bias in favor of any modality. Thirdly, students needed to fill in an online 10-item knowledge test based on a MM, and an online 10-item knowledge test based on SI. Both tests were included to study the influence of assessment modality. Figure 2 presents examples of the latter tests.

All test items were checked by two anatomical experts to guarantee equivalent difficulty over the three assessment modalities.

Covariates

Prior to data analysis, possible interacting covariates were identified. Previous research has shown that learning from 3DCBMs is strongly dependent on visuospatial ability, with significant disadvantages for low spatial ability learners.[8,10,11,23–25] Therefore, spatial ability was measured using the validated 24-item mental rotation test (MRT) described by Peters and colleagues.[26] The MRT has been validated to measure spatial ability in the general population,[26] and has an established value in anatomical education studies in the context of 3D learning.[23,24,27] Students had 6 minutes in total to complete the test. Each MRT test was scored with a maximum of 24 points.

Secondly, since the delivery modality could have an influence on mental effort, this was identified as a possible interacting covariate. The mental effort was measured after the learning intervention using the subjective Paas rating scale (9-point scale, 1 = very, very low mental effort to 9 = very, very high mental effort).[28] The scale inquired about the experienced mental effort during the learning intervention.

Data analysis

Analysis was performed with the statistic software SPSS™ 27.0 (SPSS Inc., Chicago, IL, USA). Results of the anatomy knowledge tests were presented as percentages (mean ± standard deviation (SD)). Pearson correlation coefficients and r squared values were calculated to assess the influence baseline anatomy knowledge and MRT on the dependent variables.

Independent sample t-tests were used to assess potential differences in students' baseline anatomical knowledge. To test the differential impact of the learning and testing modality on anatomy knowledge (= first and second hypothesis), one-way ANCOVA was conducted with the test scores of the physical model, SI and MM tests as dependent variables, the research conditions as fixed factors, and the pretest score as a covariate. In a second step, the MRT scores were added as a covariate.

The third hypothesis was tested with a one-way ANCOVA, using the mental effort scores as a dependent variable, the research conditions as fixed factors, and the pretest score as a covariate. In a second step, the MRT scores were added as covariate.

Finally, Cohen's d effect sizes were calculated for each comparison.

Results

The demographic data are provided in Table 1. A total of 106 students (23.4%) participated in this study. A post-hoc power analysis was done with G*Power Software (version 3.1.9.2) (minimal sample size = 84, with $\alpha = 0.05$, $\beta = 0.95$),[29] confirming the adequacy of the current sample size ($n = 106$). Sixty-three percent of the participants were female and the mean age was 18.4 (± 1.8) years. These numbers were representative for the larger course.

Table 2 gives an overview of the descriptive statistics, as well as the Pearson correlation (r) and coefficients of determination (r^2) indicating the amount of variability in posttest scores shared by baseline anatomy knowledge and MRT. Graphical representations for the interaction of the three posttest modalities with baseline anatomy knowledge and MRT are shown in Figure 3.

Baseline

The mean baseline score reflecting anatomy knowledge was 49.8% (± 0.25) and 51.3% (± 0.26) for the SI group and MM group, respectively. The independent t-test did not reveal any differences between both groups at baseline (t = -0.30, 95% CI -11.35%-8.33%, p = 0.76, Cohen's d = 0.06).

Hypothesis 1: Effect of learning modality

The mean scores and standard deviation for the two intervention groups are shown in Figure 4. The one-way ANCOVA with the pre-test scores as covariate revealed no difference among the SI and MM group (F(1, 105) = 0.06, 95% CI -10.05%-7.78%, p = 0.80, Cohen's d = 0.07). These analyses were repeated using the MRT together with the pre-test scores as covariates. The MRT did not significantly predict the scores on the test (Table 2) and the lack of differences between both groups remained (F(1,105) = 0.08, 95% CI -10.21%-7.68%, p = 0.78).

Hypothesis 2: Effect of the testing modality

a. Test with SI

   The mean scores for the two intervention groups are shown in Figure 5. The one-way ANCOVA with the pre-test scores as covariate revealed a difference among both groups (F(1, 105) = 4.21, 95% CI 0.25%-14.74%, p = 0.043, Cohen's d = 0.25). The SI group scored significantly higher than the MM group (resp. 64.2% (± 0.25) vs 57.7% (± 0.28). These analyses were repeated using the MRT together with the pre-test scores as covariates. The MRT did not significantly predict the scores on the test (Table 2) and the differences between both groups remained (F(1,105) = 4.09, 95% CI 0.14%-14.7%, p = 0.046).

b. Test with MM

The mean scores for the two intervention groups are shown in Figure 6. The one-way ANCOVA with the pre-test scores as covariate revealed no difference among the SI and MM group ($F(1, 105) = 0.27$, 95% CI -9.5%-5.56%, $p = 0.61$, Cohen's $d = 0.11$). These analyses were repeated using the MRT as covariate, together with the pre-test scores. The MRT significantly predicted the scores on the test (Table 2) but no differences between both groups could be detected ($F(1,105) = 0.39$, 95% CI -9.69%-5.03%, $p = 0.53$).

Hypothesis 3: Effect of cognitive load

No difference in reported mental effort could be demonstrated between both groups using one-way ANCOVA with the pre-test scores as covariate ($F(1, 105) = 0.64$, 95% CI -0.37-0.88, $p = 0.43$, Cohen's $d = 0.17$). This analysis was repeated with MRT as additional covariate. The MRT did not significantly predict the mental effort scores (Table 2) and the lack of any difference remained ($F(1,105) = 0.57$, 95% CI -0.39-0.86, $p = 0.45$).

DISCUSSION

Over the past few years, many institutions have made significant efforts in the integration of virtual learning resources into their anatomy curricula leading to the development of new and exciting digital anatomy models. While there is no shortage of development of these models, empirical evidence relating to their efficacy is scarce. Many of these resources were incorporated into anatomy curricula based on student perceptions.[30–32] However, their incorporation should not be solely based on student feedback. To this end, research has been done towards the effectiveness of 3DCBMs. Many researchers have set up pre-posttest designs and comparisons with no-intervention groups. However, these studies are not meaningful because “*if you teach them, they will learn*”.[33] Therefore, research should be comparative. Moreover, regarding the study of computer-based learning, it is recommended to execute effectiveness evaluations within the same level of instructional design.[19] Accordingly, the presented study aimed to address this shortfall by comparing anatomy learning based on a manipulable pelvic hip bone model versus static atlas-like images of the same model.

Hypothesis 1: Effect of learning modality

Despite principles along the embodied learning theory, test scores did not reveal a significant difference between both groups for the test on the physical bone. Literature systematically comparing multiple-view/manipulable models versus static images is very scarce and heterogeneous. However, several earlier studies found similar results. A series of experiments with multiple-view models of the carpal bones found no advantages for 3D reconstructions or interactivity over static key views.[8–10] Similarly, a study of Levinson and colleagues evaluating the effectiveness of a virtual brain model concluded that high degrees of learner control may reduce effectiveness of learning.[11] Regarding fully manipulable models, Khot and colleagues found no difference between students that learned with key views (photos of a physical model) versus a virtual model (3D reconstruction of CT scan) of the pelvis.[12] Finally, a virtual equine foot model (3D reconstruction of MRI scan) did not show to be superior over textbooks.[13] In contrast, other studies comparing monoscopic 3D to standard 2D material have determined a significant advantage, including the inner ear,[34] brain,[35] gallbladder, celiac trunk and superior mesenteric artery,[36] and the liver.[37] However, many of the studies cited above violated the research design principles recommended by Cook, comparing intervention groups with no-intervention groups, or assessing outcomes from different intervention modalities.[19,33]

A possible explanation for the counter-intuitive results in the presented study might be found in the nature of the anatomical model. It is believed that the added-value of virtual 3D models lies in the possibility to illustrate complex spatial anatomical relationships, for example the anatomy of the middle and inner ear.[34] In contrast, identification of bony structures on a pelvic bone only require basic anatomical knowledge and visuospatial skills. These results might be generalizable to all osteology models, since it was assumed that the pelvic bone is representative for other bones in the human body, and previous studies reporting no advantages of 3D reconstructions were also based on osteology models.[8–10,12,13] To sum up, it is believed that computer-based manipulable osteology models might not be superior over 2D static osteology models due to their limited spatial complexity.

Furthermore, the desktop setting might have influenced the presented results. Manipulation of 3D models in a desktop setting might not be optimally aligned with embodied cognition principles. Computer mouse manipulations are less intuitive compared to manipulating objects in the real world or immersive environments. Therefore, embodied representations might be less strong than anticipated.

Finally, Wainman and colleagues have shown that true stereoscopy might be a more important factor than active engagement.[22,38]

Hypothesis 2: Effect of testing modality

a. Test on SI

For the test on SI, the group that studied with the SI scored significantly better compared to the MM group (p = 0.043). This was anticipated because of principles along the CLT.

Based on this theory, it is believed that the memory becomes very effective when dealing with familiar material.[39] Contrarily, the MM group was not familiarized with the 2D nature of the images and the limited field of view, and therefore, they were probably handicapped compared to the SI group.

b. Test on MM

For the test on the MM, no differences between both groups could be demonstrated. However, a significant interaction with spatial ability was detected ($r = 0.18$, $p = 0.049$), demonstrating that the MRT scores appear to be a stronger predictor of success than instructional modality. The results might be explained by the presentation of the models: the SI are presented from an optimal point of view (= key views), while the MM requires rotation/manipulation to identify anatomical structures. The latter might rely heavily on spatial ability, which is confirmed by past research showing that spatial ability is an important predictor of success in 3D learning environments.[10–12,40]

Another explanation might be the limited experiential base of learning with 3DCBMs. Traditional anatomy materials at Ghent University are largely based on atlas images, prosections and dissections. Therefore, students at our university are not used to learning in 3D virtual environments. Related to the fact that high spatial ability learners experience less difficulty with the manipulation of 3DCBMs,[8–11,23,25] they might be favored in this new learning environment, whereas low spatial ability learners might have been handicapped. To assess this, future research should focus on long-term outcomes in which students can rely on a profound experiential base in manipulating 3DCBMs.

Based on these results, educators must be aware they might rather measure spatial ability than anatomy knowledge. Does this mean online 3D assessments have no place in anatomy education? Research shows that spatial ability might be trained by learning (spatial) anatomy,[27] and therefore, it is believed that 3D assessments might have a place in a more advanced stage of the anatomy curriculum.

Hypothesis 3: Effect on mental effort

Despite the hypothesis that mental effort might be reduced by giving users control over dynamic visualizations,[6] the presented study could not show any difference between both intervention groups. The extraneous load inherent to the learning modalities might explain the results: although the 3D visualizations provide the illusion of depth created by rotation and shadows, they still lack the stereoscopic sense of space provided by a physical model. In that way, the SI as well as the MM remained 2D in nature and students still needed to devote cognitive resources to inferring the 3D structure from a series of 2D images.

Finally, the MRT scores did not predict the reported mental effort. The lack of correlation might be related to the fact that students did not perceive the learning of pelvic bone anatomy as spatially complex, or the actionable affordances may not have invoked complex manipulations of the MM (e.g., panning, rotation, zooming),[41] and therefore, they do not report a heavy mental effort. In addition, based on perception studies, students tend to indicate that 3DCBMs improve their understanding of anatomical relationships and their ability to name major anatomical structures.[30,32]

Limitations

This study is subject to several limitations. First, the instruction time of 15 minutes was substantially shorter than a "real world" learning task. In addition, participants in the MM group had to spend more time clicking, panning, rotating, and zooming the model compared to the SI group. Therefore, a timespan of 15 minutes for both groups might have been unfair at the expense of the MM group.

Secondly, the pelvic bone model, and by extension osteology models, do not invoke complex spatial orientation and comprehension. Because the added value of 3D visualizations lie in the ability to display complex anatomical relationships, learning with a more spatially complex anatomy model might benefit from more sophisticated learning materials. Therefore, these results are not generalizable to other non-bony anatomical regions.

Thirdly, the outcome measurement (orientation and nominal questions) is merely a surrogate measure of 3D anatomical understanding. Posttests assessing profound spatial anatomy knowledge might be able to show advantages of 3DCBMs.

Finally, student perceptions were not measured in this study. However, since novel study materials could increase students' level of engagement and interest in anatomy, this might drive greater usage of the material, and thereby cause increased learning gains.

CONCLUSION

The mixed results in the field of 3DCBMs suggest that more systematic research is needed to guide integration of effective digital models for teaching and learning anatomy. This study is one of the few to assess the effect of active manipulation of a high-fidelity 3DCBM within the same level of instructional design. Despite principles along the embodied cognition theory, the presented study could not demonstrate any significant differences in learning performance when studying with static images versus a manipulable pelvic bone model. These counter-intuitive results suggest that certain anatomical systems may be better suited to user control while other anatomical regions do not require this as the majority of spatial and perceptual information can be expressed in key views. It is believed that for the pelvic bone, and by extension osteology models, the added-value of 3D representations is limited. Therefore, a reasonable degree of skepticism is needed when considering the integration of newer technologies, as not all educational topics might equally benefit from these technological advancements. Moreover, regarding the testing environment, the presented results show that assessing students with an online manipulable model, caution must be adopted since educators might rather evaluate spatial ability than anatomy knowledge.

Literature cited

1. Brenton H, Hernandez J, Bello F, et al. Using multimedia and Web3D to enhance anatomy teaching. *Comput Educ*. 2007;49(1):32-53. doi:10.1016/j.compedu.2005.06.005
2. Evans DJR, Bay BH, Wilson TD, Smith CF, Lachman N, Pawlina W. Going virtual to support anatomy education: a STOPGAP in the midst of the Covid-19 pandemic. *Anat Sci Educ*. 2020;13(3):279-283.
3. Harmon DJ, Attardi SM, Barremkala M, et al. An analysis of anatomy education before and during Covid-19: May–August 2020. *Anat Sci Educ*. 2021;14(2):132-147.
4. Vandenbossche V, Van de Velde J, Avet S, et al. Digital body preservation: Technique and applications. *Anat Sci Educ*. 2022;15(4):731-744.
5. Erolin C. Interactive 3D digital models for anatomy and medical education. In: P. M. Rea, ed. *Biomedical Visualisation, Advances in Experimental Medicine and Biology*. Vol 1138. Springer Nature Switzerland AG; 2019:1-16.
6. Birbara NS, Pather N. Instructional design of virtual learning resources for anatomy education. *Adv Exp Med Biol*. 2021;1317:75-110.
7. Yammine K, Violato C. A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy. *Anat Sci Educ*. 2015;8(6):525-538.
8. Garg A, Norman G, Spero L, Maheshwari P. Do virtual computer models hinder anatomy learning? *Acad Med*. 1999;74(10):S87-S89.
9. Garg A, Norman G, Sperotable L. How medical students learn spatial anatomy. *Lancet*. 2001;357(9253):363.

10. Garg A, Norman G, Eva K, Spero L, Sharan S. Is there any real virtue of virtual reality? The minor role of multiple orientations in learning anatomy from computers. *Acad Med*. 2002;77(10 SUPPL.):97-99.

11. Levinson AJ, Weaver B, Garside S, McGinn H, Norman GR. Virtual reality and brain anatomy: A randomised trial of e-learning instructional designs. *Med Educ*. 2007;41:495-501.

12. Khot Z, Quinlan K, Norman GR, Wainman B. The relative effectiveness of computer-based and traditional resources for education in anatomy. *Anat Sci Educ*. 2013;6(4):211-215.

13. Preece D, Williams SB, Lam R, Weller R. "Let's Get Physical": Advantages of a physical model over 3D computer models and textbooks in learning imaging anatomy. *Anat Sci Educ*. 2013;6(4):216-224.

14. Hackett M, Proctor M. The effect of autostereoscopic holograms on anatomical knowledge: a randomised trial. *Med Educ*. 2018;52:1147-1155.

15. Shepard RN, Metzler J. Mental rotation of three-dimensionsal objects. *Science (80-)*. 1971;171(3972):701-703.

16. Shepard RN, Cooper LA. Mental images and their transformations. *J Hist Behav Sci*. 1982;19(3):301-302.

17. Kosslyn SM, Thompson WL, Wraga M, Alpert NM. Imagining rotation by endogenous versus exogenous forces: Distinct neural mechanisms. *Neuroreport*. 2001;12(11):2519-2525.

18. Amorim MA, Isableu B, Jarraya M. Embodied spatial transformations: "body analogy" for the mental rotation of objects. *J Exp Psychol Gen*. 2006;135(3):327-347.

19. Cook DA. The research we still are not doing: An agenda for the study of computer-

based learning. *Acad Med*. 2005;80(6):541-548.

20. Bogomolova K, Sam AH, Misky AT, et al. Development of a virtual three-dimensional assessment scenario for anatomical education. *Anat Sci Educ*. 2021;9:1-9.

21. Choudhury B, Freemont A. Assessment of anatomical knowledge: Approaches taken by higher education institutions. *Clin Anat*. 2017;30(3):290-299. doi:10.1002/ca.22835

22. Wainman B, Wolak L, Pukas G, Zheng E, Norman GR. The superiority of three-dimensional physical models to two-dimensional computer presentations in anatomy learning. *Med Educ*. 2018;52:1138-1146.

23. Guillot A, Champely S, Batier C, Thiriet P, Collet C. Relationship between spatial abilities, mental rotation and functional anatomy learning. *Adv Heal Sci Educ*. 2007;12(4):491-507.

24. Lufler RS, Zumwalt AC, Romney CA, Hoagland TM. Effect of visual-spatial ability on medical students' performance in a gross anatomy course. *Anat Sci Educ*. 2012;5(1):3-9.

25. Nguyen N, Nelson AJ, Wilson TD. Computer visualizations: Factors that influence spatial anatomy comprehension. *Anat Sci Educ*. 2012;5(2):98-108.

26. Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C. A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain Cogn*. 1995;28:39-58.

27. Vorstenbosch MATM, Klaassen TPFM, Donders ART, Kooloos JGM, Bolhuis SM, Laan RFJM. Learning anatomy enhances spatial ability. *Anat Sci Educ*. 2013;6(4):257-262.

28. Paas FG, Van Merriënboer JJ, Adam JJ. Measurement of cognitive load in instructional research. *Percept Mot Skills*. 1994;79:419-430.

29. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: A flexible statistical power

analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175-191.

30. Thomas KJ, Denham BE, Dinolfo JD. Perceptions among occupational and physical therapy students of a nontraditional methodology for teaching laboratory gross anatomy. *Anat Sci Educ*. 2011;4(2):71-77.

31. Battulga B, Konishi T, Tamura Y, Moriguchi H. The effectiveness of an interactive 3-dimensional computer graphics model for medical education. *J Med Internet Res*. 2012;14(2):1-12.

32. Yao WC, Regone RM, Huyhn N, Butler EB, Takashima M. Three-dimensional sinus imaging as an adjunct to two-dimensional imaging to accelerate education and improve spatial orientation. *Laryngoscope*. 2014;124:596-601.

33. Cook DA. If you teach them, they will learn: Why medical education needs comparative effectiveness research. *Adv Heal Sci Educ*. 2012;17:305-310.

34. Nicholson DT, Chalk C, Funnell WRJ, Daniel SJ. Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. *Med Educ*. 2006;40:1081-1087.

35. Estevez ME, Lindgren KA, Bergethon PR. A Novel Three-Dimensional Tool for Teaching Human Neuroanatomy. *Anat Sci Educ*. 2010;3(6):309-317.

36. Foo JL, Martinez-Escobar M, Juhnke B, et al. Evaluating mental workload of two-dimensional and three-dimensional visualization for anatomical structure localization. *J Laparoendosc Adv Surg Tech*. 2013;23(1):65-70.

37. Müller-Stich BP, Löb N, Wald D, et al. Regular three-dimensional presentations improve in the identification of surgical liver anatomy - A randomized study. *BMC Med Educ*. 2013;13(1):131.

38. Wainman B, Pukas G, Wolak L, Mohanraj S, Lamb J, Norman GR. The critical role of stereopsis in virtual and mixed reality learning environments. *Anat Sci Educ*. 2020;13:401-412.

39. Khalil MK, Paas F, Johnson TE, Payer AF. Interactive and dynamic visualizations in teaching and learning of anatomy: A cognitive load perspective. *Anat Rec - Part B New Anat*. 2005;286:8-14.

40. Huk T. Who benefits from learning with 3D models? The case of spatial ability. *J Comput Assist Learn*. 2006;22(6):392-404.

41. Gibson JJ. The theory of affordances. In: Shaw R, Bransford's J, eds. *Perceiving, Acting, and Knowing.* Lawrence Erlbaum Associates Publishers; 1977:67-82.