

Assessing Performance and Usability of 3D Visualization Technologies for Anatomical Training

**Andrew Wismer, Lauren Reinerman-Jones, Grace Teo,
Sasha Willis, Kelsey McCracken**
University of Central Florida
Institute for Simulation and Training (UCF IST)
Orlando, FL
awismer@ist.ucf.edu, lreiner@ist.ucf.edu, gteo@ist.ucf.edu,
swillis@ist.ucf.edu, kmccrack@ist.ucf.edu

Matthew Hackett
Army Research Laboratory
Orlando, FL
matthew.g.hackett.civ@mail.mil

ABSTRACT

Anatomical structures can be displayed for educational or training purposes through several different visualization technologies. However, these technologies differ regarding their usability and display of spatial information. The present research compared three visualization technologies in the learning of gross human brain anatomy. Participants, aged 18-28 ($N = 90$, $Mdn = 18$, $IQR = 1$), were assigned to one of three conditions: (1) a PowerPoint condition, where 2D brain images with monocular depth cues were presented on a computer screen, (2) a Physical Model condition, involving a plastic 3D brain model that could be disassembled into eight parts, and (3) a Virtual Reality condition, where 3D brain models were presented using HTC Vive technology. The visualization technologies were evaluated on how well they enabled identification and spatial learning of selected brain structures, on overall usability, and on the amount of study time participants used with the technology. Participants were given a computerized test assessing their knowledge of the brain structures before and after studying with the technology. Participants in the PowerPoint condition used significantly less study time and completed questions on brain structure identification and spatial relationships faster than participants in the Physical Model and Virtual Reality conditions. There were no significant differences in identification/spatial anatomy accuracy, confidence, or technology usability. These findings provide insight into the relative costs and benefits of different 3D visualization technologies for learning gross anatomy. Future research should investigate other measures of effectiveness and learner reactions regarding the use of 3D visualization technologies in anatomical training.

ABOUT THE AUTHORS

Andrew Wismer, M.A., is a Human Factors & Cognitive Psychology Ph.D. Candidate at the University of Central Florida with research experience in categorization, decision making, anatomical training, and learning in dynamic environments.

Lauren Reinerman-Jones, Ph.D., is the Director of Prodigy, which is one lab at the University of Central Florida's Institute for Simulation and Training, focusing on assessment for explaining, predicting, and improving human performance and systems.

Grace Teo, Ph.D., is a Research Associate at the University of Central Florida's Institute for Simulation and Training. Her research includes work on decision making, individual differences, human-technology interactions, and various assessments of human performance.

Sasha Willis, B.S., is a Research Assistant at the University of Central Florida's Institute for Simulation and Training with an interest in game-based learning and three-dimensional imaging technologies.

Kelsey McCracken is an Undergraduate Research Assistant at the University of Central Florida's Institute for Simulation and Training and is currently pursuing higher education in a STEM background.

Matthew Hackett, M.S., is a Science and Technology Manager with the US Army Research Laboratory Medical Simulation and Performance Branch, conducting research in a variety of technologies to improve military medical training.

Assessing Performance and Usability of 3D Visualization Technologies for Anatomical Training

**Andrew Wismer, Lauren Reinerman-Jones, Grace Teo,
Sasha Willis, Kelsey McCracken
University of Central Florida
Institute for Simulation and Training (UCF IST)
Orlando, FL
awismer@ist.ucf.edu, lreiner@ist.ucf.edu, gteo@ist.ucf.edu,
swillis@ist.ucf.edu, kmccrack@ist.ucf.edu**

**Matthew Hackett
Army Research Laboratory
Orlando, FL
matthew.g.hackett.civ@mail.mil**

INTRODUCTION

The study of anatomy requires a learner to grasp abstract and complex spatial relationships among bodily structures. A firm understanding of gross/topographic anatomy is essential for more advanced medical and surgical training. Yet, visualizing the three-dimensional nature of anatomical structures is a difficult task for novice students and first time learners (Cottam, 1999; Dev et al., 2002; Garg, Norman, & Sperotable, 2001; Miller, 2016). Anatomy training has traditionally been imparted through lectures, textbooks, dissection/prosection labs, and the use of physical models made of fiberglass, plastic, clay, or similar material. However, the majority of learning has been from two-dimensional presentations of anatomical structures as found in lecture slides or textbooks. Learning from these formats can require a high degree of mental effort and may leave gaps in spatial understanding (Yammine & Violato, 2015; Hackett, 2013). As evidence, some researchers have found medical students are not on par with practical standards regarding anatomical knowledge (Cottam, 1999; Smith, 2005; Waterston & Stewart, 2005; Aung, Htet, San, & Fen, 2000).

Anatomy instruction and training is evolving to meet the needs forgone by conventional methods through advances in visualization technologies (Skochelek, 2010). There are a variety of three-dimensional visualization technologies (3DVTs) available today to support the acquisition of anatomical knowledge in medical education. Many researchers have provided evidence of the benefits of three-dimensional training platforms compared to more traditional 2D presentations (Müller-Stich et al., 2013; Ruisoto, Juanes, Contador, Mayoral, & Prats- Galino, 2012; Silén, Wirell, Kvist, Nylander, & Smedby, 2008; Petersson, Sinkvist, Wang, & Smedby, 2009; Hilbelink, 2009; Hackett & Proctor, 2016). However, few studies have provided a direct comparison of multiple types of 3DVTs (e.g., Lombardi, Hicks, Thompson, & Marbach-Ad, 2014). A critical comparison of different 3DVTs may provide a better understanding of when, where, and for whom these technologies might best be applied. A comparison of performance and usability among respective 3DVTs will provide guidance for the effective implementation of these technologies into the anatomy curriculum. In view of recent cutbacks in curricular hours in anatomical instruction (see Papa & Vaccarezza, 2013), these technologies will be valuable as a supplementary method to traditional training methods (i.e., cadaveric dissection; Habbal, 2009; Codd & Choudhury, 2011; Ghosh, 2017).

The present work evaluates three 3DVTs ranging in methods of displaying spatial information and level of interaction afforded: a 3D monoscopic display, a physical model, and 3D models presented in virtual reality. Three-dimensional monoscopic displays provide spatial information through the use of monocular depth cues, such as shading, elevation, interposition, and absolute size (see Schwartz, 2010). This type of display provides a relatively simple way of displaying 3D information; however, its presentation on a flat 2D surface limits its immersiveness and interaction capabilities (Prajapati, Madrigal, & Friedman, 2016). Physical models, on the other hand, have become more standard in anatomical instruction (Baskaran, Strkalj, Strkalj, & Di Ieva, 2016; McMenamin, Quayle, McHenry, & Adams, 2014). Physical models are realistic replicas of an anatomical structure that allow the learner to feel, rotate, and disassemble the model to aid development of mental representations through increased depth information and haptic feedback. Finally, virtual reality provides a digital alternative to physical models, immersing the learner in a virtual world where the learner can interact with the 3D models in a way similar to that with the physical model. However, a distinct advantage with virtual reality models is that they allow models to be updated relatively quickly and the learner can have access to libraries of 3D anatomical models, such as through the Human Anatomy Atlas (2016).

The effectiveness of different visualization technologies for anatomical knowledge gain is not readily determined as some studies show advantages for physical models (Preece, Williams, Lam, & Weller, 2013; Qamar, Ahmad, & Ashar,

2014), while others report on the superiority of virtual models (Brewer, Wilson, Eagleson, & de Ribaupierre, 2012; Khot, Quinlan, Norman, & Wainman, 2013), and yet others show no differences between the two compared to traditional anatomical training with images presented in a two-dimensional plane on a textbook page (Codd & Choudhury, 2011; Keedy et al., 2011). The present work evaluated the usability of three different 3DVTs and their efficacy for acquiring spatial knowledge of anatomical structures. It was expected that there would be differences among the three visualization technology conditions, although the expected pattern of differences was unclear.

METHODS

Participants

Students from the University of Central Florida comprised the study sample. All participants were at least 18 years of age, reported normal or corrected-to-normal vision, and were of no vulnerable population. The data set is comprised of ninety participants (Males = 46 [51%], Females = 44 [49%]), with ages ranging from 18 to 28 years (*Mdn.* = 18, *IQR* = 1). Each participant provided written informed consent, completed the experiment within three hours, and received course credit for their participation. Participants were similar in age, gender, educational background, and number of anatomy courses taken (see Table 1).

Table 1. Average Demographics by 3DVT Condition and Overall

Measure	PowerPoint <i>n</i> = 29	Physical Model <i>n</i> = 29	Virtual Reality <i>n</i> = 32	Total <i>N</i> = 90	
Age (<i>Mdn</i> and <i>IQR</i>)	<i>Mdn</i> = 18.00 <i>IQR</i> = 2	<i>Mdn</i> = 18.00 <i>IQR</i> = 2	<i>Mdn</i> = 18.00 <i>IQR</i> = 1	<i>Mdn</i> = 18.00 <i>IQR</i> = 1	
Gender (Count and %)	Males = 17 (58.6%) Females = 12 (41.4%)	Males = 16 (55.2%) Females = 13 (44.8%)	Males = 13 (40.6%) Females = 19 (59.4%)	Males = 46 (51.1%) Females = 44 (48.9%)	
Highest Education Level Completed (Count and %)	High School:	24 (82.8%)	19 (65.5%)	28 (87.5%)	71 (78.9%)
	2 Years of College:	4 (13.8%)	7 (24.1%)	4 (12.5%)	15 (16.7%)
	Bachelor's Degree:	1 (3.4%)	3 (10.3%)	0 (0.0%)	4 (4.4%)
Number of Anatomy Courses (Count and %)	0 = 19 (65.5%)	0 = 19 (65.5%)	0 = 20 (62.5%)	0 = 58 (64.4%)	
	1 = 10 (34.5%)	1 = 5 (17.2%)	1 = 9 (28.1%)	1 = 24 (26.7%)	
	2 = 0 (0.0%)	2 = 4 (13.8%)	2 = 3 (9.4%)	2 = 7 (7.8%)	
	3 = 0 (0.0%)	3 = 1 (3.4%)	3 = 0 (0.0%)	3 = 1 (1.1%)	

Note. Median age is reported here due to positive skew. Mean age overall was 19.09 (2.20); however, most participants (62%) were 18 years old. The remaining measures are reported as counts and percentages.

Experimental Design

The present study adopted a 3 (3D visualization technologies) x 2 (Pre-task vs. Post-task) mixed ANOVA design with the 3D visualization technologies as the between-subjects factor. The study sought to compare the effectiveness of three different 3D visualization technologies (3DVTs) in knowledge acquisition of human brain anatomy during a timed learning task. The three 3DVT conditions included a 3D monoscopic display in the form of PowerPoint (PPT) slides, a Physical Model (PM), and 3D Virtual Reality (VR) models.

Materials

Participants were asked to study the location of and spatial relationships among 16 macroscopic brain structures using the 3DVT to which they were assigned. Although every effort was made to ensure that the images and views across

the 3DVTs were comparable, there were some capabilities unique to each technology, as the study aim was to evaluate the capabilities of each technology in helping learners acquire spatial information.

Monoscopic 3D PowerPoint

In the PowerPoint (PPT) condition, participants were presented three slides on a standard computer monitor, with each slide containing a separate key view of the human brain (i.e., lateral, ventral, and medial views). In this condition, the researcher indicated to the participant that they may use either the keyboard arrows or the mouse to switch between the slides. The key views were of 3D monoscopic images (i.e., images that look identical whether viewed by one or two eyes; see Schwartz, 2010) as seen from a lateral, ventral, and a medial perspective of the brain. For each image, the structures of interest visible in that view were labeled (see Figure 1). The researchers explained to participants that because the brain is symmetrical, structures were only labeled once per slide but could be found on both sides of the brain and could also be found on other slides from a different angle. The images shown in the PPT condition were snapshots of the brain models used in the Virtual Reality condition and were appropriately matched to the brain model used in the Physical Model condition. The images used for this condition displayed monocular 3D depth cues such as shading, texture gradients, and linear perspective.

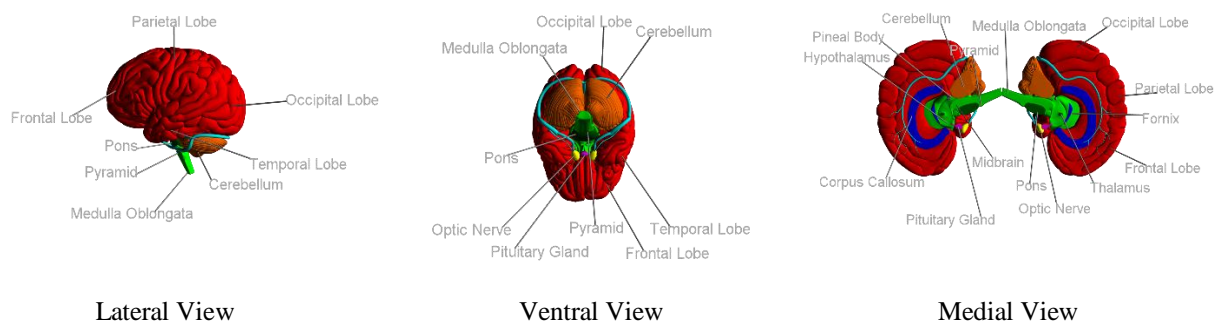


Figure 1. Three monoscopic, key views of the brain models that were displayed in the PPT condition

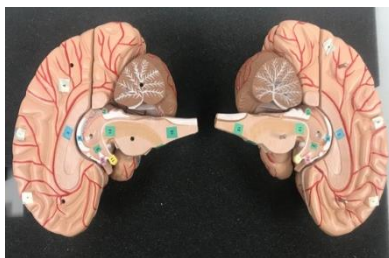


Figure 2. Physical Model (Axis Scientific)

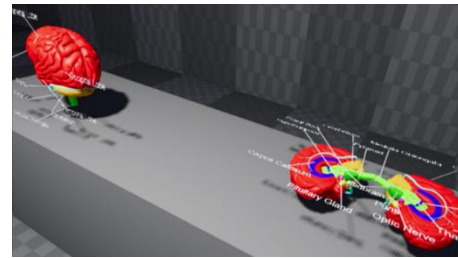
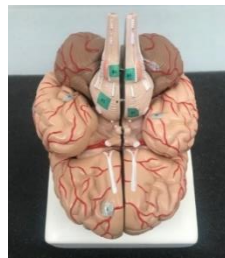


Figure 3. Virtual Reality Models

Physical Model

Participants in the Physical Model (PM) condition received a replica of the human brain made of high-quality PVC plastic by Axis Scientific, measuring 6 x 5.5 x 5.5 inches and weighing 2.5 pounds. The model consisted of eight separate pieces, held together by small metal rods and weak magnets, which allowed participants to view the brain as a whole or in disassembled pieces (see Figure 2). In this way, the viewer could observe the same lateral, ventral, and medial views as the other conditions but could also examine the internal structures of the brain from additional viewpoints. The 16 brain structures of interest were numbered on the plastic model and the key which mapped the numbers to the full names of the structures was placed next to the model. The researcher pulled apart each of the pieces to demonstrate how the participant could choose to use the technology, while describing the labels and their matching structure identifier on the key.

Virtual Reality

The Virtual Reality (VR) condition displayed two, computer-generated models of the human brain using the HTC Vive. One model displayed an external view of the brain with two sets of labels: one for viewing a lateral view of the model and one for a ventral view. The second model displayed the brain from a medial plane (see Figure 3). The label

positions matched placement in the PPT condition. Furthermore, the virtual model allowed participants to rotate the brain, revealing a 360-degree view of the external architecture, which was afforded by the PM condition as well but not the PPT condition. One difference between the VR condition and the other conditions lay in the ability of participants to toggle “on” (structure name visible) and “off” (structure name removed) the labels on the models, a feature that was only available with this computer generated technology.

Due to the novelty of VR technology compared to the other two technologies, the researcher presented a more in-depth explanation of the technology by guiding the participant through a training scenario. The training scenario was identical to the learning task except the two brain models were replaced with colored cubes with external rings representing the structure labels. The training scenario was used to teach the participant the buttons, task, and interaction capabilities to ensure the participant was comfortable to complete the learning task on their own.

Measures

The three 3DVTs were evaluated on the following criteria: (i) the level of spatial knowledge they allowed participants to acquire, (ii) level of confidence in knowledge acquired, (iii) time taken to answer spatial anatomy questions, (iv) study time needed for participants to acquire that knowledge, and (v) the level of usability of the 3DVT.

Level of Spatial Knowledge Acquired

A Spatial Anatomy Test (SAT) was developed to assess the level of spatial knowledge acquired using the 3DVTs. The 35-question SAT presented three different types of questions: identification (16 questions), spatial multiple choice (15 questions), and mental rotation (4 questions). Identification questions required participants to locate and identify one of the 16 structures of interest (i.e., “name structure labelled ‘A’ on the diagram”). Multiple choice questions, such as those used by Hilbelink (2009), required identification of the names of structures and an understanding of where they were located in relation to other structures; in other words, these questions gauged a participant’s understanding of spatial relationships. An example would be a question asking which brain structure was nearest a given structure. Mental rotation questions, similar to that developed by Nguyen, Mulla, Nelson, and Wilson (2014), required no knowledge of structure names, nor did participants have to see the structures before being able to answer these types of questions; however, a spatial understanding of the brain may help in answering these questions. For example, participants were asked which two of four rotated images matched a target image (with two images being mirrored copies). A 40-second time limit also accompanied mental-rotation type questions, deemed appropriate by a pilot study.

The SAT was administered to participants prior to and subsequent to experiencing the learning task, designated as pre-task SAT and post-task SAT, respectively. The purpose of the pre-task SAT was to determine how much knowledge of human brain anatomy participants had prior to experiencing the learning task, independent from self-ratings of knowledge on anatomy. The post-task SAT was identical to the pre-task SAT except in the presentation order of questions. Participants were not told of their score after taking the pre-task SAT, so as not to influence their learning or subsequent performance on the post-task test.

Confidence

The level of confidence in knowledge acquired through the 3DVT was also assessed during the SAT. A confidence rating question followed each of the four mental rotation questions with an additional two questions at the end of the test: one for overall confidence in identification questions and one for overall confidence in multiple-choice questions. Confidence ratings were on a scale from 1 (very low) to 5 (very high).

SAT Completion Time

The time taken to complete the SAT was a measure of the efficacy of the 3DVTs for spatial learning of brain anatomy. Completion times for each SAT question type (i.e., identification, multiple-choice, mental rotation) were recorded.

Study Time

Another aspect of efficacy of the 3DVTs in anatomical training was the duration it took for participants to acquire the spatial knowledge from each technology. Although participants were limited to ten minutes in the learning task, they were given the option to end early if they felt confident. During the PPT and PM conditions, a timer counting down from 10 minutes was displayed on a second monitor in front of the participant. In the VR condition, this timer was displayed on the back wall of the virtual environment. Once the participant notified the researcher they were satisfied

with the amount of time they spent studying, or the timer ran out, the researcher ended the scenario and prepared the participant for the next portion of the experiment.

Usability

The System Usability Scale (SUS; Brook, 1986) was administered to participants following the learning task and post-task SAT. The SUS contains 10 items that assessed a participant's subjective experience of a technology. Each statement concerning usability (e.g., "I felt very confident using the system.") was accompanied by a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). While overall usability ratings can be evaluated for each 3DVT experienced by participants, individual items from the SUS may provide additional insight into the different effectiveness of the technologies.

Procedure

After obtaining their informed consent, participants were first instructed to complete pre-task surveys. This comprised a demographics survey and a pre-task Spatial Anatomy Test (SAT) that included an item where participants self-rated their pre-existing knowledge of the anatomy. After these were completed, participants were introduced to their assigned 3DVT and were instructed on its use. Participants were then given the opportunity to familiarize themselves with technology until they were able to use the technology on their own. They were subsequently given up to ten minutes to study the location and relationship of 16 brain structures using the 3DVT. Participants were able to end the task early if they felt confident in their ability to recall the relevant material. They were then tested on the knowledge acquired with a post-task SAT, and they rated the usability of the 3DVT on the System Usability Scale (SUS). After the study session, participants were thanked, debriefed, and dismissed.

RESULTS

To determine the effect of 3D Visualization Technology (3DVT) on various measures, 3 x 2 mixed factor ANOVAs were conducted with 3D Visualization Technology (3DVT) as the between-subjects variable and testing (pre-task SAT, post-task SAT) as the within-subjects variable. Bonferroni's method was used to assess pairwise comparisons to decrease the possibility of a Type I error. Due to the use of a pre- and post-task test design, the interaction results of the ANOVAs will be reported rather than the main effects as they provide information on differences in knowledge gain among 3DVT conditions.

Pre-Task Differences

Pre-task differences between the conditions were evaluated with a series of one-way between subjects ANOVAs. There were no differences in the self-ratings of pre-existing anatomy knowledge across the conditions, $F(2, 87) = 1.29, p = .279, \eta_p^2 = .029$. This concurred with the results of the one-way between-subjects ANOVA conducted on pre-task SAT overall accuracy scores, which indicated there were no significant differences in pre-task spatial knowledge of brain anatomy among the 3DVT conditions, $F(2, 87) = 0.90, p = .410, \eta_p^2 = .020$, nor were there differences for multiple choice accuracy, $F(2, 87) = 0.37, p = .692, \eta_p^2 = .008$. There was a significant difference among conditions, however, for identification accuracy, $F(2, 87) = 3.47, p = .035, \eta_p^2 = .074$, and mental rotation accuracy, $F(2, 87) = 3.25, p = .044, \eta_p^2 = .069$. There were no differences in the pre-task SAT confidence ratings among the 3DVT conditions for identification, $F(2, 87) = 2.48, p = .090, \eta_p^2 = .054$, or multiple choice questions, $F(2, 87) = 1.31, p = .274, \eta_p^2 = .029$. However, the conditions differed on mental rotation confidence ratings, $F(2, 87) = 4.21, p = .018, \eta_p^2 = .088$. Participants in the PM condition ($M = 3.16, SD = 0.62$) had higher mental rotation confidence ratings in the pre-task SAT than both the PPT ($M = 3.16, SD = 0.62; p = .036$) and VR ($M = 3.20, SD = 0.89; p = .046$) conditions. In general, differences among the conditions prior to the study treatment were non-existent or negligible.

Level of Spatial Knowledge Acquired

Mixed factor ANOVAs were conducted on average overall, identification, multiple choice, and mental rotation accuracy scores to evaluate differences in SAT accuracy gain among 3DVT conditions (see Figure 4). There was no interaction between 3DVT and testing on overall accuracy scores, $F(2, 87) = 0.39, p = .682, \eta_p^2 = .009$, identification accuracy scores, $F(2, 87) = 0.46, p = .631, \eta_p^2 = .011$, multiple choice accuracy scores, $F(2, 87) = 0.14, p = .874, \eta_p^2 = .004$.

= .871, or mental rotation accuracy scores, $F(2, 87) = 0.31, p = .737, \eta_p^2 = .007$. This suggests that each 3DVT condition affords a relatively equal spatial knowledge gain in novice learners.

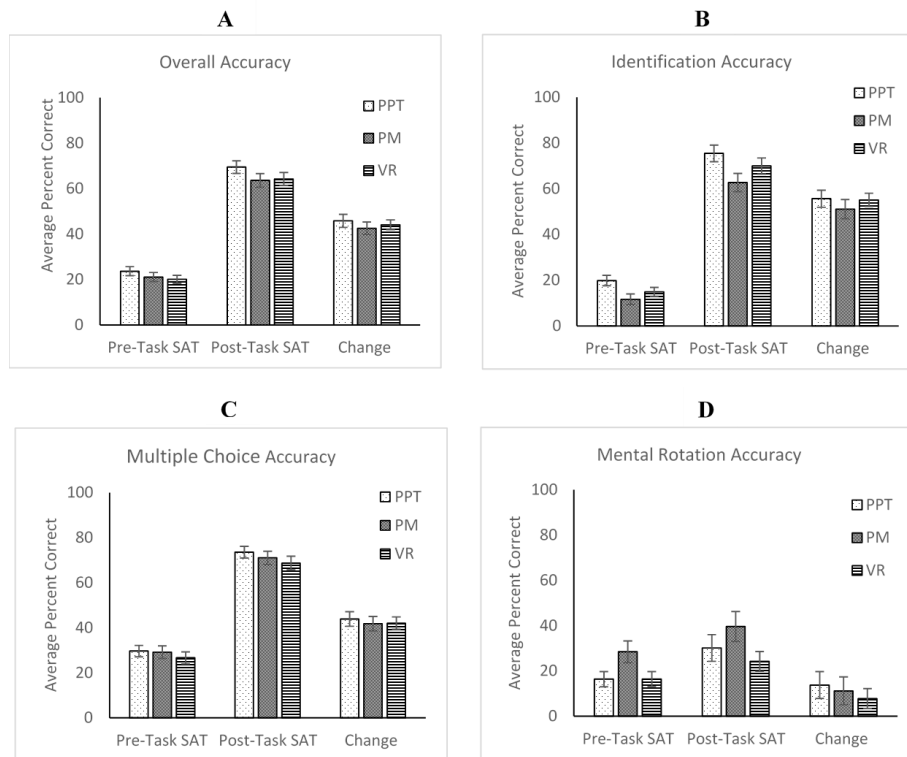


Figure 4. Average overall (A), identification (B), multiple choice (C), and mental rotation (D) accuracy on SAT by 3DVT condition.

Confidence

Mixed factor ANOVAs were then conducted on identification, multiple choice, and average mental rotation confidence ratings to evaluate differences in the change in SAT completion time from pre-task to post-task among 3DVT conditions (see Table 2). There were no significant interactions between 3DVT and testing on identification, $F(2, 87) = 0.41, p = .667, \eta_p^2 = .009$, multiple choice, $F(2, 87) = 0.43, p = .654, \eta_p^2 = .010$, or mental rotation confidence ratings, $F(2, 87) = 0.59, p = .554, \eta_p^2 = .013$.

Table 2. Confidence Ratings on SAT by Question Type and 3DVT Condition

Question Type	Condition	Pre-Task SAT <i>M (SD)</i>	Post-Task SAT <i>M (SD)</i>	Change <i>M (SD)</i>
Identification	PowerPoint	2.00 (1.20)	4.03 (1.02)	2.03 (1.57)
	Physical Model	1.41 (0.78)	3.62 (0.98)	2.21 (1.29)
	Virtual Reality	1.72 (0.99)	3.59 (0.95)	1.88 (1.43)
Multiple Choice	PowerPoint	1.59 (0.91)	3.66 (1.08)	2.03 (1.10)
	Physical Model	1.34 (0.61)	3.48 (1.06)	2.21 (1.30)
	Virtual Reality	1.66 (0.79)	3.53 (0.92)	1.88 (1.07)
Mental Rotation	PowerPoint	3.16 (0.62)	3.51 (0.78)	0.34 (0.70)
	Physical Model	3.70 (0.83)	3.86 (0.92)	0.16 (0.59)
	Virtual Reality	3.20 (0.89)	3.49 (0.83)	0.30 (0.67)

Note. Confidence ratings were on a Likert scale from 1 (very low) to 5 (very high).

SAT Completion Times

Mixed factor ANOVAs were conducted on average overall, identification, multiple choice, and mental rotation completion times to evaluate differences in the change in completion time from pre-task and post-task SAT among 3DVT conditions (see Table 3). There was an interaction between 3DVT and testing on overall completion times, $F(2, 87) = 4.90, p = .010, \eta_p^2 = .101$. Participants in the PM condition took significantly longer ($M = 488.59, SD = 141.33$) to complete the post-task SAT than participants in the PPT condition ($M = 403.26, SD = 73.79; p = .013$)¹, took longer to complete the post-task than pre-task SAT ($p < .001$)², and had a greater change in overall completion time ($M = 88.27, SD = 125.11$) than both participants in the PPT ($M = -4.35, SD = 109.20; p = .003$)³ and VR ($M = 19.41, SD = 117.06; p = .024$)⁴ conditions. There was also an interaction between 3DVT and testing for identification completion times, $F(2, 87) = 3.64, p = .030, \eta_p^2 = .077$, such that post-task identification completion times were significant for participants in the PM condition ($M = 193.47, SD = 59.53$) than both PPT ($M = 133.96, SD = 32.55; p < .001$)⁵ and VR ($M = 157.97, SD = 49.66; p = .016$)⁶ conditions. Additionally, there was an interaction between 3DVT and testing for multiple choice completion times, $F(2, 87) = 5.05, p = .008, \eta_p^2 = .104$, such that all 3DVT conditions took longer to complete multiple choice questions during the post-task SAT than they did to complete the same questions during the pre-task SAT, but the increase in time was greatest for the PM condition^{7,8}. There was no significant interaction between 3DVT and testing on mental rotation completion times, $F(2, 87) = 1.42, p = .247, \eta_p^2 = .032$.

Table 3. SAT Completion Times (in seconds) by Question Type and 3DVT Condition

Question Type	Condition	Pre-Task SAT <i>M (SD)</i>	Post-Task SAT <i>M (SD)</i>	Change <i>M (SD)</i>
Overall	PowerPoint	407.62 (93.53)	403.26 (73.79) ¹	-4.35 (109.20) ³
	Physical Model	400.32 (151.73) ²	488.59 (141.33) ^{1, 2}	88.27 (125.11) ^{3, 4}
	Virtual Reality	428.54 (121.99)	447.95 (108.27)	19.41 (117.06) ⁴
Identification	PowerPoint	159.93 (47.38)	133.96 (32.55) ⁵	-25.97 (14.83)
	Physical Model	167.51 (82.51)	193.47 (59.53) ^{5, 6}	25.96 (22.98)
	Virtual Reality	173.86 (65.43)	157.97 (49.66) ⁶	-15.90 (15.76)
Multiple Choice	PowerPoint	144.23 (40.62)	181.40 (41.73)	37.16 (-1.11) ⁷
	Physical Model	134.82 (67.05)	215.22 (95.72)	80.39 (-28.67) ^{7, 8}
	Virtual Reality	155.90 (58.02)	198.90 (62.02)	43.00 (-4.00) ⁸
Mental Rotation	PowerPoint	103.46 (25.37)	87.91 (32.22)	-15.55 (-6.85)
	Physical Model	97.99 (29.93)	79.90 (29.45)	-18.09 (0.47)
	Virtual Reality	98.77 (23.12)	91.08 (29.64)	-7.69 (-6.52)

Note: Dashed borders indicate a significant effect. Superscript numbers signify significant post-hoc pairwise comparisons. Refer to the SAT Completion Times results section for an interpretation of the differences.

Study Time

Participants were given 10 minutes to study 16 brain structures during the learning task; however, they could elect to end their study time early if they felt confident enough to continue before the 10 minutes had elapsed. Learning task study times are reported in seconds (maximum 600 seconds). A one-way between-subjects ANOVA was conducted to analyze the study times among the 3DVT conditions, and a significant difference in study time was found, $F(2, 87) = 14.40, p < .001, \eta_p^2 = .249$. Pairwise comparisons indicated there was a significant difference in study time during the learning task between participants in the PPT condition ($M = 419.97, SD = 155.99$) and both the VR ($M = 568.53, SD = 59.45; p < .001$) and PM conditions ($M = 541.84, SD = 109.84; p < .001$). These results show participants in the PPT condition studied for a significantly shorter amount of time than did participants in the other two conditions. There was no significant difference in study time between participants in the VR and PM conditions ($p > .999$).

Usability

A one-way between-subjects ANOVA was conducted to evaluate differences in usability among 3DVT conditions. There was no main effect of 3DVT on usability ratings, $F(2, 87) = 1.06, p = .351, \eta_p^2 = .024$. On a scale from 0 to 100,

usability ratings were fairly high for all conditions: PPT ($M = 75.66$, $SD = 14.21$), PM ($M = 79.10$, $SD = 12.36$), and VR ($M = 81.00$, $SD = 16.30$).

There was, however, a significant difference found among conditions for the first question on the SUS. This question asked participants to rate the statement, “I think that I would like to use this system frequently”. A one-way between-subjects ANOVA revealed a significant difference in expected frequency of system use among 3DVT conditions, $F(2, 87) = 6.29$, $p = .003$, $\eta_p^2 = .126$. Pairwise comparisons indicated there was a significant difference in expected frequency of system use between participants in the VR ($M = 4.06$, $SD = 1.13$) and PPT conditions ($M = 3.03$, $SD = 1.24$; $p = .002$). These results show that participants in the VR condition would like to use the system (i.e., Vive headset) more frequently than participants in the PPT condition would like to use the PPT system. There were no significant differences between participants in the VR and PM conditions ($p = .616$) or between participants in the PPT and PM conditions ($p = .094$).

DISCUSSION

This study compared three 3-dimensional visualization technologies (3DVTs; PowerPoint [PPT], Physical Model [PM], Virtual Reality [VR]) on their ability to support knowledge gain in spatial anatomy in novice learners as well as for overall usability for anatomical training. There were no significant differences in the change in accuracy or confidence among 3DVT conditions, nor were there any differences in perceived usability. However, participants in the PPT condition studied for a shorter amount of time than both the PM and VR conditions, and their reduction in completion time on the Spatial Anatomy Test (SAT) from pre- to post-task was greater.

Learning anatomy involves the acquisition and retention of factual and spatial aspects of anatomical structures. The present work addresses the acquisition component of anatomical training in novice learners. In anatomical knowledge acquisition, it is certainly important to learn structure locations and relationships. It is also important to acquire the knowledge in a limited amount of time, given the time constraints in traditional anatomy courses (Papa & Vaccarezza, 2013; Ghosh, 2017). The present work suggests that, at least for gross brain anatomy, a variety of 3DVTs may enable successful acquisition of spatial anatomical information. However, more subtle differences in study time and test completion times provide information on the efficacy of one 3DVT versus others.

Despite being limited to a short ten-minute study period with the 3DVT, participants who studied images on PowerPoint slides spent less time studying, on average, than participants studying with either a physical model or virtual reality. Therefore, for the present task, PowerPoint images (i.e., a 3D monoscopic display) enabled individuals to learn spatial anatomical information as well as those who used physical or virtual models, with less time required for studying. Additionally, participants in the PPT condition decreased the amount of time they took to complete a test on spatial brain anatomy to a significantly greater extent than participants in PM or VR conditions. Thus the speed of learning, or time to confidence, was greater in the PPT condition.

These results suggest a fundamental advantage for learning gross spatial anatomy using 3D monoscopic displays – which are the least complex type of three-dimensional visualization technology. However, a few arguments could be made against the generalization of this finding. First, PowerPoint is a common presentation tool used in classrooms. It is possible that familiarity with PowerPoint and the simplicity of the three key views resulted in shorter study and test times. Differences in study time might be explained by differences in learning curve and interaction affordability, as the PM and VR conditions involved a higher degree of interaction to study the brain structures. Findings from Jang, Vitale, Jyung, and Black (2017) suggest that direct manipulation during learning most benefits people with low spatial ability. It is possible then that our sample as a whole had relatively high spatial ability and thus learned successfully with the PPT key views. Future research should evaluate the potential moderating role of spatial ability in 3DVTs.

Still, previous research comparing key views to interactive conditions similarly showed no advantage for VR over key views (Garg, Norman, Spero, & Maheshwari, 1999; Khot et al., 2013; Levinson, Weaver, Garside, McGinn, & Norman, 2007). In addition, Garg, Norman, Eva, Spero, & Sharan (2002) found no difference in performance, after controlling for spatial ability, between a multi-view condition (MV) where the learner controlled the viewing angle and a key views + wiggle room condition (KV+W) where the learner could see only key views plus a small degree of rotation around these views. The researchers found participants in the MV condition spent most of their time studying

the key viewpoints, even though they had access to a full range of orientations. As a whole, these studies suggest that the PPT condition in the present study may have benefited from presenting limited, key views of the brain.

A second argument against generalization of the present findings is that participants were tested on images presented on the computer screen on which the PowerPoint slides were also presented. It is possible that the higher fidelity between the learning and testing conditions for those in the PPT condition resulted in the improved performance in terms of study and test times. Previous research by Metzler et al. (2012) provides evidence for this proposition, as learning with 3D models did not improve testing on 2D images. This is a valid argument that should be addressed in future work. Although the present work did display 2D test images on a computer screen, the test images were matched to the model type seen during learning.

The advantages seen in the PPT condition do not negate the potential value of physical and virtual reality models, however. Both physical and virtual models allow for learner interaction and tactile engagement. They allow for learner control and viewing multiple perspectives to facilitate three-dimensional understanding. Given the ability for learner interaction, physical and virtual models provide the possibility for procedural training. Virtual models can be obtained from online image repositories allowing for quick model updates and repeated practice. The benefits of these model types must be weighed against any known costs (monetary or otherwise) in selecting the appropriate visualization technology. Thus, while some advantages were seen in the present study for 3D monoscopic PowerPoint images, caution must be taken before generalizing to a host of other applications. Physical and/or virtual models may provide learning benefits for other anatomical structures, for knowledge retention, or for procedural skills.

Finally, future work should consider both how individual differences (such as mental rotation ability, personality, or previous experience) influence knowledge gain with 3DVTs, as well as the effect of an individual 3DVT on other learner reactions (such as engagement, enjoyment, or workload). These additional sources of information will enable a more complete picture of the utility of three-dimensional visualization technologies for anatomical training. One final important component to study would be the retention of spatial anatomical information. In particular, it is possible that learner reactions to a 3DVT may influence long term retention of the studied information. Nonetheless, the work presented here suggests there is value in providing 3D monoscopic images, such as on PowerPoint slides, to aid novice learners in acquiring spatial anatomical information.

ACKNOWLEDGEMENTS

This research was sponsored by the U.S. Army Research Laboratory (ARL) and was accomplished under Cooperative Agreement Number W911NF-15-2-0011. The views and conclusions contained in this document are those of the author's and should not be interpreted as representing the official policies, either expressed or implied, of ARL or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation herein.

REFERENCES

- Aung, H. H., Htet, H., San, G. P., & Fen, S. C. C. (2000). Anatomical knowledge among Medical, Dental and Health Science Professionals.
- Baskaran, V., Štrkalj, G., Štrkalj, M., & Di Ieva, A. (2016). Current applications and future perspectives of the use of 3D printing in anatomical training and neurosurgery. *Frontiers in neuroanatomy, 10*.
- Brewer, D. N., Wilson, T. D., Eagleson, R., & de Ribaupierre, S. (2012). Evaluation of neuroanatomical training using a 3D visual reality model. In *MMVR* (pp. 85-91).
- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry, 189*(194), 4-7.
- Codd, A. M., & Choudhury, B. (2011). Virtual reality anatomy: is it comparable with traditional methods in the teaching of human forearm musculoskeletal anatomy?. *Anatomical sciences education, 4*(3), 119-125.
- Cottam, W. W. (1999). Adequacy of medical school gross anatomy education as perceived by certain postgraduate residency programs and anatomy course directors. *Clinical Anatomy, 12*(1), 55-65.
- Dev, P., Montgomery, K., Senger, S., Heinrichs, W. L., Srivastava, S., & Waldron, K. (2002). Simulated medical learning environments on the Internet. *Journal of the American Medical Informatics Association, 9*(5), 437-447.
- Garg, A., Norman, G. R., Spero, L., & Maheshwari, P. (1999). Do virtual computer models hinder computer learning? *Academic Medicine, 74*(Suppl 10), S87-S89.

- Garg, A. X., Norman, G. R., Eva, K. W., Spero, L., & Sharan, S. (2002). Is there any real virtue of virtual reality?: The minor role of multiple orientations in learning anatomy from computers. *Academic Medicine*, 77(10), S97-S99.
- Garg, A. X., Norman, G., & Sperotable, L. (2001). How medical students learn spatial anatomy. *The Lancet*, 357(9253), 363-364.
- Ghosh, S. K. (2017). Cadaveric dissection as an educational tool for anatomical sciences in the 21st century. *Anatomical sciences education*, 10(3), 286-299.
- Habbal, O. (2009). The state of human anatomy teaching in the medical schools of Gulf Cooperation Council countries: Present and future perspectives. *Sultan Qaboos University Medical Journal*, 9(1), 24.
- Hackett, M. (2013). Medical holography for basic anatomy training. ARMY RESEARCH LAB ORLANDO FL.
- Hackett, M., & Proctor, M. (2016). Three-dimensional display technologies for anatomical education: a literature Review. *Journal of Science Education and Technology*, 25(4), 641-654.
- Hilbelink, A. J. (2009). A measure of the effectiveness of incorporating 3D human anatomy into an online undergraduate laboratory. *British Journal of Educational Technology*, 40(4), 664-672.
- Human Anatomy Atlas. (2016). Visual Body. Accessible at: <https://www.visiblebody.com/anatomy-and-physiology-apps/human-anatomy-atlas>
- Jang, S., Vitale, J. M., Jyung, R. W., & Black, J. B. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computers & Education*, 106, 150-165.
- Keedy, A. W., Durack, J. C., Sandhu, P., Chen, E. M., O'Sullivan, P. S., & Breiman, R. S. (2011). Comparison of traditional methods with 3D computer models in the instruction of hepatobiliary anatomy. *Anatomical sciences education*, 4(2), 84-91.
- Khot, Z., Quinlan, K., Norman, G. R., & Wainman, B. (2013). The relative effectiveness of computer- based and traditional resources for education in anatomy. *Anatomical sciences education*, 6(4), 211-215.
- Levinson, A. J., Weaver, B., Garside, S., McGinn, H., & Norman, G. R. (2007). Virtual reality and brain anatomy: a randomized trial of e-learning instructional designs. *Medical education*, 41(5), 495-501.
- Lombardi, S. A., Hicks, R. E., Thompson, K. V., & Marbach-Ad, G. (2014). Are all hands-on activities equally effective? Effect of using plastic models, organ dissections, and virtual dissections on student learning and perceptions. *Advances in physiology education*, 38(1), 80-86.
- McMenamin, P. G., Quayle, M. R., McHenry, C. R., & Adams, J. W. (2014). The production of anatomical teaching resources using three- dimensional (3D) printing technology. *Anatomical sciences education*, 7(6), 479-486.
- Metzler, R., Stein, D., Tetzlaff, R., Bruckner, T., Meinzer, H. P., Büchler, M. W., ... & Fischer, L. (2012). Teaching on three-dimensional presentation does not improve the understanding of according CT images: a randomized controlled study. *Teaching and learning in medicine*, 24(2), 140-148.
- Miller, M. (2016). Use of computer-aided holographic models improves performance in a cadaver dissection-based course in gross anatomy. *Clinical Anatomy*, 29, 917-924.
- Müller-Stich, B. P., Löb, N., Wald, D., Bruckner, T., Meinzer, H. P., Kadmon, M., ... & Fischer, L. (2013). Regular three-dimensional presentations improve in the identification of surgical liver anatomy—a randomized study. *BMC medical education*, 13(1), 131.
- Nguyen, N., Mulla, A., Nelson, A. J., & Wilson, T. D. (2014). Visuospatial anatomy comprehension: The role of spatial visualization ability and problem- solving strategies. *Anatomical sciences education*, 7(4), 280-288.
- Papa, V., & Vaccarezza, M. (2013). Teaching anatomy in the XXI century: new aspects and pitfalls. *The Scientific World Journal*, 2013.
- Petersson, H., Sinkvist, D., Wang, C., & Smedby, Ö. (2009). Web- based interactive 3D visualization as a tool for improved anatomy learning. *Anatomical sciences education*, 2(2), 61-68.
- Prajapati, S., Madrigal, E., & Friedman, M. T. (2016). Acquisition, Visualization and Potential Applications of 3D Data in Anatomic Pathology. *Discoveries*, 4(4), 1-12.
- Preece, D., Williams, S. B., Lam, R., & Weller, R. (2013). “Let's get physical”: advantages of a physical model over 3D computer models and textbooks in learning imaging anatomy. *Anatomical sciences education*, 6(4), 216-224.
- Qamar, K., Ahmad, A., & Ashar, A. (2014). Comparison of learning anatomy with cadaveric dissection and plastic models by medical students. *Pakistan Armed Forces Medical Journal*, 64(2).
- Ruisoto, P., Juanes, J. A., Contador, I., Mayoral, P., & Prats- Galino, A. (2012). Experimental evidence for improved neuroimaging interpretation using three- dimensional graphic models. *Anatomical Sciences Education*, 5(3), 132-137.
- Schwartz, S. (2010). *Visual perception: A clinical orientation* (4th ed.). McGraw Hill Professional.
- Silén, C., Wirell, S., Kvist, J., Nylander, E., & Smedby, Ö. (2008). Advanced 3D visualization in student-centred medical education. *Medical teacher*, 30(5), e115-e124.
- Skochelak, S. E. (2010). A decade of reports calling for change in medical education: what do they say?. *Academic Medicine*, 85(9), S26-S33.
- Smith, J. A. (2005). Can anatomy teaching make a come back?. *ANZ journal of surgery*, 75(3), 93-93.
- Waterston, S. W., & Stewart, I. J. (2005). Survey of clinicians' attitudes to the anatomical teaching and knowledge of medical students. *Clinical Anatomy*, 18(5), 380-384.
- Yamine, K., & Violato, C. (2015). A meta- analysis of the educational effectiveness of three- dimensional visualization technologies in teaching anatomy. *Anatomical sciences education*, 8(6), 525-538.