# Computer-based Learning of Neuroanatomy: A Longitudinal Study of Learning, Transfer, and Retention

**Julia H. Chariker (julia.chariker@louisville.edu)**
**Farah Naaz (farah.naaz@louisville.edu)**
**John R. Pani (jrpani@louisville.edu)**
Department of Psychological and Brain Sciences, University of Louisville
Louisville, KY 40292 USA

**Abstract**

Using interactive computer-based methods of instruction, this research examined the contribution of whole (3D) anatomical knowledge to learning sectional anatomy. Participants either learned sectional anatomy alone or learned whole anatomy prior to learning sectional anatomy. Sectional anatomy was explored either with perceptually continuous navigation or discretely, as in the use of an anatomical atlas. Learning occurred over repeated cycles of study, test, and feedback, and continued to a high performance criterion. After learning, transfer of knowledge to interpreting biomedical images and long-term retention were tested. Whole anatomy was learned quickly and transferred well to the learning of sectional anatomy: initial accuracy was higher, learning of sectional anatomy was completed more rapidly, and there was less error over the entire course of learning. Knowledge of whole anatomy benefited the long-term retention of sectional anatomy at 2-3 weeks. Learners demonstrated high levels of transfer to the interpretation of biomedical images.

**Keywords:** learning; transfer; computer; anatomy.

## Introduction

In medicine and many areas of science, anatomy education serves as a vital foundation for high level knowledge and skill. Unfortunately, anatomy is challenging to learn. Large volumes of material must be learned in relatively short periods of time. Anatomical structures often have irregular and indistinct shapes. They have little variation in color and texture, and they are related to each other in complex three-dimensional arrangements. Moreover, a comprehensive education in anatomy extends to include a thorough knowledge of sectional anatomy, which is necessary for diagnostic imaging, microscopy, and dissection.

Sectional anatomy is particularly challenging to learn. A spatial transformation occurs when a two-dimensional section is taken from a three-dimensional object. The two and three-dimensional structures may look very different from each other. In addition, multiple mappings are possible between these representations of anatomy. One-to-many mappings occur because anatomy can be sectioned at different depths and orientations, resulting in significant variation in the presentation of structures across a series of sections. Many-to-one mappings occur because differently shaped structures can appear similar in a sectional image.

The challenges in learning sectional anatomy might be reduced by facilitating cognitive organization of the mass of information in the sections (consider Bower, Clark, Lesgold, & Winzenz, 1969). Given that anatomical sections are derived from whole anatomy, helping students develop a thorough understanding of the shapes and relationships of whole structures prior to learning sectional anatomy would seem an ideal way to help students organize the information in the sections. The benefit of organization for learning and memory has been established for verbal materials, but it is not clear what effect organization has in domains where spatial reasoning is required.

Knowledge of whole anatomy may also serve as a mental model that supports reasoning about sectional anatomy. Reasoning has been found to play a large role in the successful interpretation of histological sections viewed under the microscope (e.g., Pani, Chariker, & Fell, 2005).

A second approach to helping students organize information in sectional anatomy may be in the presentation of sectional anatomy itself. Serial presentation of the sections would be expected at a minimum, but additional support may be found by providing smooth, seamless navigation through the sections. Work in *anorthoscopic perception* and *kinetic completion* suggests that with this approach, learners may see the series of sections as a unified whole. On the other hand, continuous presentation of sectional anatomy can be considered a form of animation, and there has been mixed success in using animation in instruction (e.g., Hegarty, 2005; Tversky, Morrison, & Betrancourt, 2002).

In the current study, we explored both approaches to organizing sectional anatomy. Half of the participants learned whole anatomy before learning sectional anatomy (*transfer* groups), while the other half learned only sectional anatomy (*sections alone* groups). Within each of these groups, half of the participants learned sectional anatomy using a continuous presentation, and half learned with a discrete presentation -- analogous to turning the pages of an anatomical atlas.

Participants learned neuroanatomy in interactive computer-based environments. This approach holds potential for helping learners build rich mental representations of anatomy. For example, a computer-based model of 3D anatomy can be rotated to allow exploration of anatomy from any angle. It can be virtually dissected, restored to its original state, and then dissected again.

The instructional programs were designed to promote efficient learning through a method that we call *adaptive exploration*. With graphical models and exploratory tools

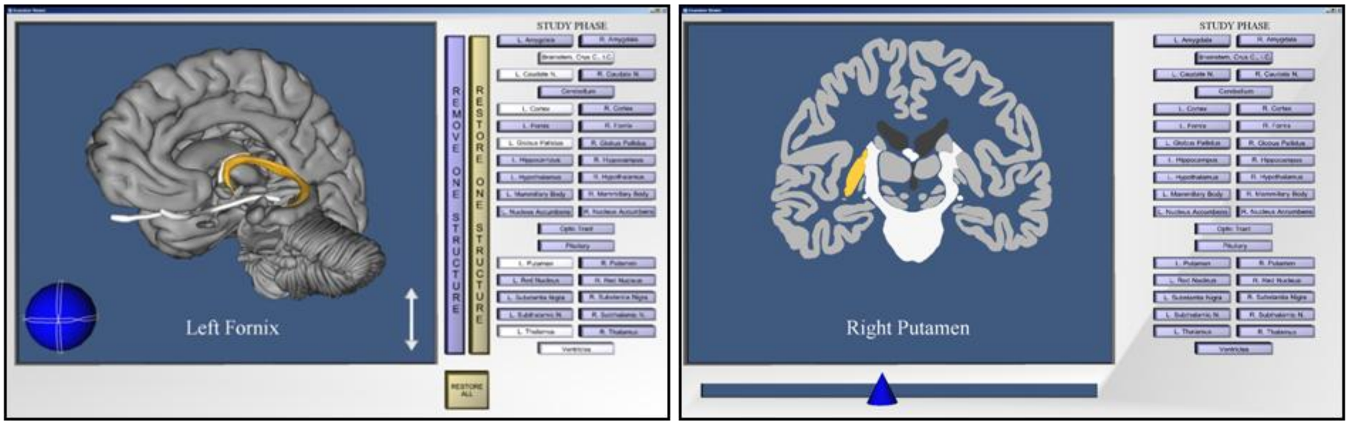


Figure 1: Screenshots of the anatomical model and the interface in the study phase of the whole anatomy learning program (left) and the sectional anatomy learning program (right).

available, learning was measured over multiple trials of study, test, and feedback until a high performance criterion was reached. In testing and feedback, participants learned the nature of the test to be mastered and were continually updated on progress in learning. This information allowed learners to adaptively adjust exploration of anatomy during study. Additionally, this approach to learning conforms to what appears to be best practices in regard to optimizing long-term retention through repeated testing (e.g., Karpicke & Roediger, 2008).

All participants learned 19 neuroanatomical structures across three standard views of anatomy: coronal, sagittal, and axial. After learning was completed, we measured the degree to which participants could transfer anatomical knowledge to interpreting biomedical images. Retention of anatomical knowledge was measured 2-3 weeks after learning was completed.

## Method

### Participants

Seventy-two undergraduate students at the University of Louisville were recruited for the study through advertisements placed around campus. All were at least 18 years of age. Only those respondents who reported minimal knowledge of neuroanatomy were enrolled. Participants were paid $8.00 per hour for their participation.

Each participant was administered the Space Relations subtest of the Differential Aptitude Tests, a test of spatial ability, prior to beginning the study (DAT-SR; Bennett, Seashore, & Wesman, 1989). The mean and the distribution of scores were balanced across the four learning groups.

### Materials

A three-dimensional (3D) computer graphical model of the human brain was created for this research (see Figure 1). Digital images of neuroanatomical cryosections in the Visible Human project (Vers. 2.0) of the National Library of Medicine were used as source material for the model (Ratiu, Hillen, Glaser, & Jenkins, 2003). The brain model is composed of 19 structures, including the cerebral cortex, ventricles, cerebellum, brainstem, amygdala, caudate nucleus, fornix, globus pallidus, hypothalamus, hippocampus, mammillary bodies, nucleus accumbens, optic tract, pituitary, putamen, red nucleus, substantia nigra, subthalamic nucleus, and thalamus. The structures were colored in dark gray (ventricles), medium gray, and white to approximate the basic appearance of light and dark structures in typical biomedical images of the brain.

Three relatively dense sets of serial sections were created from the brain model. There were 60 coronal sections, 50 sagittal sections, and 46 axial sections. All sections were taken at equal intervals.

MRI images were used to test transfer of knowledge. The images were made available from the SPL-PNL Brain Atlas (Kikinis et al., 1996). The images are typical gray scale T1 images of structures in the head and neck. The images were slightly brightened and contrast enhanced and presented at a screen resolution of 895 x 895 pixels. Visible Human images also were used to test transfer. These images were from the Visible Human 2.0 dataset. The images were high resolution color images of structures in the head and neck.

Computer programs for learning neuroanatomy were created using the C++ programming language and the Open Inventor library for interactive graphics. There was a common format for all of the learning programs. The differences between the programs were modifications related to the type of anatomy presented and the different presentations of sectional anatomy.

In all of the learning programs, a participant completed two learning trials -- one block of trials -- before a single run of the program terminated. Participants were presented with the same form and view of anatomy (e.g., sectional anatomy, coronal view) throughout the two trials in a block.

Each learning trial was composed of three phases: study, test, and feedback. Throughout each phase, tools were available that functioned specifically for either whole or sectional anatomy. In the study phase, participants had three minutes to freely explore the brain. On selecting a structure, its name appeared on the screen. In the test phase, the participant's task was to identify the anatomical structures in the model. Testing was self-paced. In the feedback phase, the participant saw the same orientation of the brain, and used the same tools and procedures, as in the study phase. In addition, structures were color coded to provide participants with information about their performance on the test.

In the study and feedback phases for whole anatomy learning, a rotation tool allowed participants to smoothly rotate the model 360 degrees forward and backward or right and left. A zoom tool allowed participants to move the model closer or further from view. Buttons were available that allowed participants to remove or restore structures.

In the test phase of a trial, model rotation was constrained to a total range of 90 degrees of motion -- 45 degrees in any direction from the initial viewpoint. This ensured that a participant's performance on the test was specific to the viewpoint being learned in that trial.

Two programs were created for learning sectional anatomy, one for the continuous and one for the discrete form of navigation. In the study phase of a trial, both programs presented a set of anatomical sections in serial order in a single viewing plane. There was a slider at the bottom of the screen, and the two learning programs differed in the way the slider functioned. In the continuous program, moving the slider resulted in continuous movement back and forth through the series of sections. A section of the brain was always visible, and the transition between sections comprised a type of animation. In addition, a highlighted structure remained highlighted in each section in which it appeared.

In the discrete presentation program, movement between sections was perceptually discontinuous. When participants moved the slider, the brain became invisible. The number of the corresponding section in the series appeared prominently at the bottom of the viewing area. On stopping at a numbered section, a 0.75 second delay occurred before the appropriate section of the brain appeared. When participants moved to a new section, highlighting was removed.

The test phase of a learning trial was the same in the two programs. Participants were given a series of test sections, presented one at a time. In each section, one or more structures were indicated with a red arrow, and the participant's task was to correctly label those structures. Although all 19 structures were tested in each trial, the section of a structure that was tested varied across trials.

During the feedback phase of the trial, participants used the slider to find each of the test sections in the series. A message reading "Test Section" appeared prominently on the screen when a test section was accessed by the slider. The tested structures in each test section were identified with the same red arrows that appeared in the test.

Three computer programs were created to test transfer of knowledge to the interpretation of biomedical images. In the first test, Uncued Recognition, participants were presented with a set of 9 images, one at a time, and asked to identify all of the structures they thought they recognized in each image. Participants identified structures by indicating the location of a structure with the mouse (leaving a red dot on the image) and then selecting the name of the structure from a list on the interface. The images alternated through coronal, sagittal, and axial views, in that order.

The remaining two test programs provided cues to the presence of structures in the images. In the Submit Structure test, the name of a single structure was presented at the bottom of each image, and participants selected the appropriate structure in the image. In the Submit Name test, a single structure was designated by a red arrow in each image, and participants selected its name from a list on the interface. Each test was comprised of three subtests, one for each view of anatomy.

A sectional anatomy test and a whole anatomy test were created for testing long-term retention. For participants who had only seen sectional anatomy, the test of whole anatomy was a test of transfer rather than retention. These tests were the same as tests given during learning and were created for all three views of anatomy.

## Apparatus

Participants sat individually at computer workstations with large high resolution LCD screens (24 inch, 1200 x 1952 pixels). Participants were tested alone in small quiet rooms with the doors closed.

## Design and Procedure

The core experimental design was a 2 X 2 between-groups factorial: anatomy course (transfer vs. sections alone) by sectional anatomy presentation (continuous vs. discrete).

Prior to beginning any of the learning or testing programs in the study, participants were trained on all aspects of the task using instructional software developed for this purpose.

During the learning portion of the study, performance in identifying 19 neuroanatomical structures was measured over multiple blocks of trials. Percent correct was calculated for each trial, and mean percent correct was calculated for each block of two trials. Participants continued learning anatomy until they reached a minimum of 89.5 percent accuracy (17 of 19 structures) in each of three consecutive learning blocks—all three views of anatomy. Across blocks of learning trials and throughout testing, the order in which view was presented was standardized at coronal, followed by sagittal, and then axial.

Immediately after learning was completed, participants were given the three tests of transfer to biomedical images in the order Uncued Recognition, Submit Structure, and Submit Name. For each test, participants were tested with each image type (MRI and Visible Human) in all three views of anatomy. The two image types were counterbalanced across participants.

Two to three weeks after learning was completed, participants were given the test of long-term retention for sectional anatomy followed by the test of long-term retention/transfer for whole anatomy. Tests were given for all three views of anatomy.

## Results

### Learning

**Learning Trajectories** Multilevel modeling was used for statistical analysis of performance in learning (Raudenbush & Bryk, 2002). Binomial models were appropriate for these data. Variables tested for inclusion in the multilevel model included learning block, anatomy course (AC), sectional anatomy presentation (SAP), and spatial ability (DAT-SR). Spatial ability was a significant factor in each of the models of learning but will not be discussed in this paper. Details of model parameters are available from the authors.

To establish the relative efficiency of learning whole anatomy and sectional anatomy, the transfer group's performance in whole anatomy was compared to the sections alone group's performance in sectional anatomy. Participants learning whole anatomy had substantially higher performance in the first block of trials and learned at a faster rate than participants learning sectional anatomy (see Figure 2). Mean percent correct identification in block one was 54 percent for whole anatomy and 36 percent for sectional anatomy, $t(69) = 5.780$, $p < .001$. Both groups improved in performance over successive blocks, $t(68) = 15.746$, $p < .001$; however, the increase in performance was much greater for participants learning whole anatomy: AC, $t(68) = 7.359$, $p < .001$.

There were no effects on the efficiency of learning sectional anatomy due to the type of sectional anatomy presentation in any of the analyses of learning. This variable was not retained in the multilevel models and will not be discussed further in the presentation of results on learning.

In a second analysis, transfer of learning from whole to sectional anatomy was measured by comparing performance in sectional anatomy for the transfer and sections alone groups. Participants in the transfer groups performed significantly better in the first block of sectional anatomy learning than participants in the sections alone groups (see Figure 2). Mean percent correct identification was 73 percent in the transfer groups and 36 percent in the sections alone groups, $t(69) = 13.522$, $p < .001$. Although both groups improved over time, the transfer groups continued their learning at a slower rate than the sections alone groups: AC, $t(70) = -3.321$, $p = .002$.

In a third analysis, differences between conditions were further explored by comparing performance in sectional anatomy for the transfer and sections alone groups after relating performance to the total time spent learning neuroanatomy. For the transfer groups, learning blocks were numbered to reflect the time participants spent learning both whole and sectional anatomy. Nearly two thirds of the participants in the Transfer groups (21 of 36) completed whole anatomy learning in 4 blocks and transferred to sectional anatomy in block 5. Therefore, performance in sectional anatomy learning was compared beginning at block 5. Modeled performance in Block 5 was 71 percent for the Transfer groups and 81 percent for the Sections alone groups, AC, $t(69) = -3.030$, $p = .004$. The 10 percent difference is equivalent to 2 of the 19 structures on the test.

**Learning Time to Achieve Criterion Performance** In each learning trial, time was constrained to 3 minutes for study and 3 minutes for feedback. Therefore, we considered the number of blocks required to reach the performance criterion as one measure of learning efficiency. An ANCOVA was performed to compare the number of blocks of trials required to complete learning for whole and

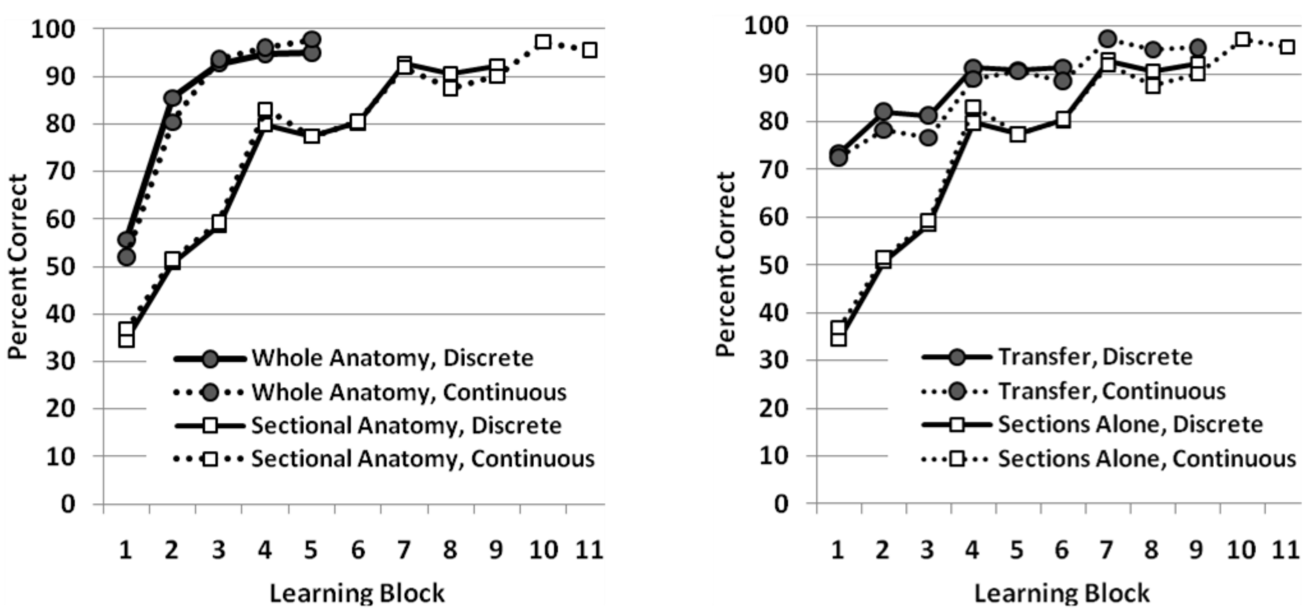


Figure 2: A comparison of performance in whole anatomy and sectional anatomy (left) and a comparison of performance in sectional anatomy beginning at block 1 (right).

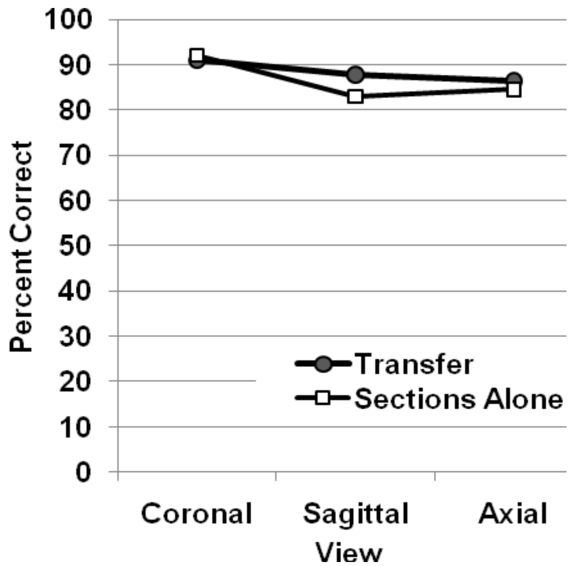


Figure 3: Sectional anatomy retention.

sectional anatomy. Spatial ability was correlated with the number of blocks required ($r$ = -.236, $p$ = .046) and was entered as a covariate, $F(1,67) = 7.785$, $p = .007$. Participants learned whole anatomy in significantly fewer blocks ($M$ = 5.2) than participants learned sectional anatomy ($M$ = 10.7), $F(1, 67) = 57.555$, $p < .001$.

A second ANCOVA compared the transfer and the sections alone groups on the number of trial blocks required to reach criterion in sectional anatomy learning. Again, spatial ability was correlated with the number of blocks to reach criterion ($r$ = -.298, $p$ = .011) and was included as a covariate, $F(1,67) = 7.678$, $p = .007$. Participants in the transfer groups completed sectional anatomy in 2.5 fewer blocks ($M$ = 8.2) than participants in the sections alone groups ($M$ = 10.7), $F(1, 67) = 7.282$, $p = .009$.

A third ANCOVA was performed to look for differences between the groups in the number of blocks of trials necessary to complete all learning in neuroanatomy. Spatial ability was correlated with the number of blocks to reach criterion ($r$ = -.344, $p$ = .003) and was included as a covariate, $F(1,67) = 10.129$, $p = .002$. Participants in the transfer groups completed whole anatomy and sectional anatomy in 2.7 more blocks than participants in the sections alone groups completed sectional anatomy (transfer, $M$ = 13.4; sections alone, $M$ = 10.7), $F(1, 67) = 6.021$, $p = .017$.

**Total Error in Learning Neuroanatomy** Over the entire course of learning, participants in the transfer groups made fewer errors ($M$ = 77) in learning neuroanatomy than participants in the sections alone groups ($M$ = 100), $F(1, 67) = 3.870$, $p = .053$. This occurred even though the transfer groups were required to complete two presentations of anatomy and took 2.7 more blocks to do so. Spatial ability was a significant covariate in the analysis of total error, $F(1, 67) = 13.995$, $p < .001$.

## Testing

**Long-Term Retention and Transfer** MANCOVA was used to analyze retention of sectional anatomy and retention/transfer of whole anatomy. DAT-SR was included as a covariate.

Retention of sectional anatomy remained high two to three weeks after learning, with several participants reaching 100% accuracy in the first test (see Figure 3). There was an interaction of AC with view, Wilks' Lambda ($\Lambda$) = .898, $F(2, 63) = 3.570$, $p = .034$. The transfer groups were more accurate than the sections alone groups for retention of the sagittal view of sectional anatomy (transfer $M$ = 87.8, sections alone $M$ = 83.1), $t(57) = -2.675$, $p = .03$ (Bonferroni). No differences between the groups occurred for retention of the coronal and axial views.

In the analysis of retention/transfer for whole anatomy, participants in the transfer groups were more accurate than participants in the sections alone groups in identifying whole anatomy, $F(1, 64) = 15.306$, $p < .001$. Participants in the transfer groups tested at 97% mean accuracy in identifying whole brain structures. Although participants in the sections alone groups had never seen whole anatomy, they reached an overall mean accuracy of 89.5%. This meets the numerical criterion used for successful learning. Given this high rate of transfer, it is important to consider that there was a relatively substantial effort required to achieve this performance. All tests in this experiment were self-paced. In an analysis of test duration, participants in the sections alone groups took substantially more time than the transfer groups to complete the three tests for whole anatomy ($M$ = 14.5 minutes vs. 8.8 minutes, a difference of nearly 6 minutes), $F(1, 61) = 54.331$, $p < .001$. This

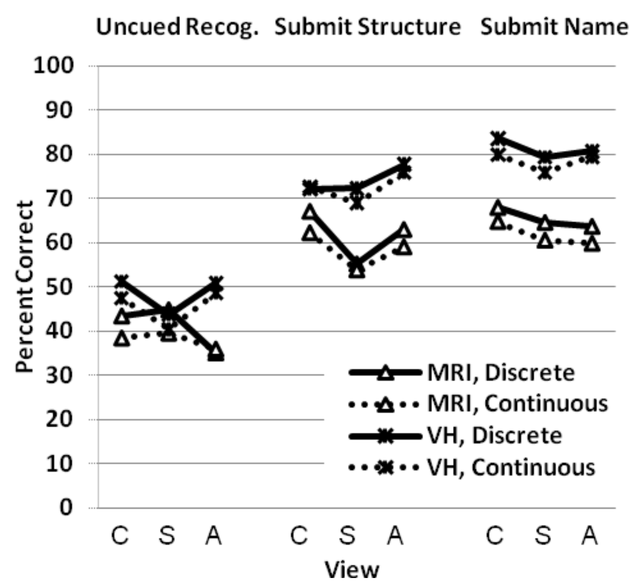


Figure 4: Transfer to MRI and Visible Human images for the discrete and the continuous sectional anatomy presentation groups.

suggests that participants who had received sectional learning alone were not recalling a representation of whole anatomy but were inferring it.

**Transfer to Biomedical Images** In scoring Uncued Recognition and Submit Structure, correct answers were decided ahead of time, and images were created with the structure boundaries drawn on them. During scoring, the experimenters were blind to the participants' identities and experimental conditions. MANOVA was used to analyze performance on each test.

Transfer performance was quite high, particularly for the two cued tests (Submit Structure and Submit Name; see Figure 4). Within each test, performance varied widely among individuals, with some participants performing extremely well. In Submit Structure and Submit Name, the best performing participants were above 90% accuracy.

There were no differences in transfer due to learning group. Performance was higher for Visible Human than for MRI images in all three tests: Uncued Recognition, VH $M$ = 47%, MRI $M$ = 40%, $\Lambda$ = .440, $F(1, 65) = 82.659$, $p < .001$; Submit Structure, VH $M$ = 72%, MRI $M$ = 58%, $\Lambda$ = .704, $F(1, 64) = 26.913$, $p < .001$; Submit Name, VH $M$ = 80%, MRI $M$ = 64%, $\Lambda$ = .378, $F(1, 64) = 105.282$, $p < .001$.

In two of the three transfer tests, Uncued Recognition and Submit Name, there was a main effect of sectional anatomy presentation: Uncued Recognition (continuous $M$ = 41.7, discrete $M$ = 44.9), $F(1, 65) = 3.962$, $p = .051$; Submit Name (continuous $M$ = 70.1, discrete $M$ = 73.4), $F(1, 64) = 4.835$, $p = .032$. Participants who learned with a discrete presentation were more accurate in identifying structures than participants who learned with a continuous presentation.

## Discussion

Knowledge of whole anatomy served as an effective basis for learning sectional anatomy. Whole anatomy was learned quickly—in half of the time of sectional anatomy. Knowledge of whole anatomy transferred well to learning sectional anatomy. Accuracy in block 1 of sectional anatomy was twice as high for the transfer groups, and learning of sectional anatomy was completed more quickly. There was less error over the entire course of learning for participants learning both representations of anatomy.

Knowledge of whole anatomy benefited long term retention of sectional anatomy. Because the participants who learned whole anatomy required *fewer* trials with sectional anatomy, this advantage for retention is inconsistent with the well-known test effect. In the test effect, a *greater* number of tests of knowledge during learning leads to an advantage for long-term retention. However, tests administered during learning and at retention are identical to each other. For the present research, such a test effect would show better long-term retention for the *sections alone* groups.

On the other hand, the groups that learned both whole and sectional anatomy did require more total trials to learn. Thus, the improvement in long-term retention is potentially due to a type of test effect, one that we have not seen described elsewhere. In this case, additional testing of whole anatomy is contributing to the long-term retention of sectional anatomy, a further instance of transfer of learning.

The transfer of knowledge to the interpretation of biomedical images served as a gold-standard test of the present methods of computer-based learning of neuroanatomy. The high levels of transfer obtained, along with the high levels of long-term retention, strongly encourage the use of these methods in neuroanatomy instruction.

## Acknowledgments

This research was supported by a grant from the National Library of Medicine, National Institutes of Health to J. Pani (Grant 1 R01 LM008323; Histological Reasoning: Visual Cognition in Microanatomy).

## References

Bennett, G.K., Seashore, H. G., & Wesman, A. G. (1989). *Differential Aptitude Tests for Personnel and Career Assessment: Space Relations.* San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovich.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior, 8,* 323-343.

Hegarty, M. (2005). Multimedia learning about physical systems. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319,* 966-968.

Kikinis, R., Shenton, M. E., Iosifescu, D. V., McCarley, R. W., Saiviroonporn, P., Hokama, H. H., … Jolesz, F. A. (1996). A digital brain atlas for surgical planning, model driven segmentation and teaching. *IEEE Transactions on Visualization and Computer Graphics, 2,* 232-241.

Pani, J. R., Chariker, J. H., & Fell, R. D. (2005). Visual cognition in microscopy. *Cogsci 2005: Proceedings of the XXVII Annual Conference of the Cognitive Science Society, 27,* 1702-1707.

Ratiu, P., Hillen, B., Glaser, J., & Jenkins, D. P. (2003). Visible Human 2.0: The next generation. In J.D.Westwood, H. M. Hoffman, G. T. Mogel, R. Phillips, R. A. Robb, & D. Stredney (Eds.), *Medicine Meets Virtual Reality 11 -- NextMed: Health Horizen*. Amsterdam, The Netherlands: IOS Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage Publications.

Tversky, B. Morrison, J. B., Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human Computer Studies, 47,* 247-262.