



Published in final edited form as:

Anat Sci Educ. 2013 ; 6(1): 11–18. doi:10.1002/ase.1297.

Computer-Based Learning: Interleaving Whole and Sectional Representation of Neuroanatomy

John R. Pani*, Julia H. Chariker, and Farah Naaz

Laboratory for Visual-Spatial Learning, Department of Psychological and Brain Sciences, College of Arts and Sciences, University of Louisville, Louisville, Kentucky

Abstract

The large volume of material to be learned in biomedical disciplines requires optimizing the efficiency of instruction. In prior work with computer-based instruction of neuroanatomy, it was relatively efficient for learners to master whole anatomy and then transfer to learning sectional anatomy. It may, however, be more efficient to continuously integrate learning of whole and sectional anatomy. A study of computer-based learning of neuroanatomy was conducted to compare a basic transfer paradigm for learning whole and sectional neuroanatomy with a method in which the two forms of representation were interleaved (alternated). For all experimental groups, interactive computer programs supported an approach to instruction called *adaptive exploration*. Each learning trial consisted of time-limited exploration of neuroanatomy, self-timed testing, and graphical feedback. The primary result of this study was that interleaved learning of whole and sectional neuroanatomy was more efficient than the basic transfer method, without cost to long-term retention or generalization of knowledge to recognizing new images (Visible Human and MRI).

Keywords

neuroanatomy education; interleaving; transfer of learning; computer assisted instruction; computer based learning; instruction; 3D graphics; generalization

INTRODUCTION

Instruction of neuroanatomy encompasses multiple subdomains and forms of representation. The curriculum includes, for example, cortical gyri and sulci, subcortical nuclei, neural pathways, white matter tracts, blood vessels, and cranial nerves. These are presented in 3D anatomy, 2D sectional images, and systematic descriptions. Such a large volume of material requires efficient instruction and extensive self-study (consider Sugand et al., 2010).

We have begun to develop and test systems for interactive computer-based instruction of neuroanatomy (Chariker et al., 2011, 2012). Such systems should be efficient to deploy, and they should be beneficial for a domain in which graphical representation is central and extended self-study is required. Many groups have seen this potential, and several systems

*Correspondence to: Dr. John R. Pani, Department of Psychological and Brain Sciences, University of Louisville, 358A Life Sciences Building, Louisville, KY 40292, USA. jrpani@louisville.edu, Phone: (502) 852-3956; Fax: (502) 852-8904.

are available or under construction (UW-SIG, 1999; Höhne et al., 2003; Crowley and Medvedeva, 2006; Luursema et al., 2006; Nicholson et al., 2006; Kockro et al., 2007; Primal Pictures, 2007; Nowinski et al., 2009; Trelease, 2008; Petersson et al., 2009; Adams and Wilson, 2011; Solomon et al., 2011; Yeung et al., 2011). However, this work is in early stages of development, and large improvements in capability can be expected (consider Issenberg et al., 2005). Related to this situation, recent reviews have suggested that additional research is needed to assess the value of computer-based instruction in biomedicine (Cook, 2005; Issenberg et al., 2005; Winkelman, 2007; Collins, 2008; Ruiz et al., 2009; Tam et al., 2009; Malone et al., 2010).

The basic method of instruction that our group has implemented so far is *adaptive exploration* (Pani et al., 2005a; Chariker et al., 2011). Software tools are provided for flexible and efficient graphical exploration of the content of a domain. These tools are used in the context of repeated trials of study, test, and feedback. Over trials, exploration is informed by an understanding of how knowledge must be used, where learning has progressed, and what is left to be learned (Mahmood and Darzi, 2004; Pashler et al., 2005; Kornell et al., 2009). When a high level of test performance is required before learning is stopped, there will be numerous learning trials, extensive testing, and spaced practice. These aspects of a learning environment are beneficial for long term retention (Cepeda et al., 2006; Karpicke and Roediger, 2008; Larsen et al., 2008; Rohrer and Pashler, 2010).

Our research program has been focused on the problem of learning sectional representation. This is a common form of representation in which 3D structures are viewed as planar slices that have been sampled from the interiors of the structures. Examples include microscope slides, tissue slices preserved in transparent media, the Visible Human dataset, and MRI and CT scans. Prior research has suggested that understanding anatomy through sectional imagery is particularly challenging and requires years to master (Lesgold et al., 1988; Crowley et al., 2003; Pani et al., 2005b; see also Ruisoto et al., 2012).

The problem of learning sectional neuroanatomy was approached from a long-standing tradition in the psychology of learning and memory that emphasizes the importance for memory of making material to be learned conceptually well organized (Bower et al., 1969; Bower, 1970; Bransford et al., 2000). An initial hypothesis was that learning whole (three-dimensional) neuroanatomy first would lead to: (1) efficient learning of sectional neuroanatomy by allowing better organized encoding of sectional geometry, (2) better long term retention of sectional neuroanatomy by providing the benefits of close integration between related forms of representation, and (3) better generalization of knowledge to the interpretation of new representations of neuroanatomy by providing a mental model that aids in reasoning.

An initial study lent strong support to this hypothesis (Chariker et al., 2011, 2012). Whole neuroanatomy was learned quickly and transferred well to learning sectional neuroanatomy. Initial learning of sectional neuroanatomy was relatively rapid, and fewer trials of sectional neuroanatomy learning were required to reach criterion. As a result, the total errors committed while learning neuroanatomy were fewer for learning both forms of representation than for learning sectional neuroanatomy alone. In addition, when both forms

of representation had been learned, there was a small benefit for the generalization of knowledge to interpreting new sectional imagery and for long-term retention of sectional neuroanatomy.

The focus on integrating two forms of representation leads to a question of whether it may be possible to improve the instructional method. Whole and sectional anatomy are alternative spatial representations of the same structures. It is possible that learning would be most effective if there was close integration during the process of learning through alternating the two forms of representation. Questions and discoveries that arose with one type of representation could be related immediately to study of the other. In addition, there would not be an increasingly long delay between the initial learning of whole anatomy and its application to the learning of sectional anatomy. The final cognitive mapping between the two domains of representation should be articulated in more detail. Prior work on the benefits of *interleaving* different types of material during learning encouraged examining this approach in a domain that featured complex spatial representation (Rohrer and Taylor, 2007; Kornell and Bjork, 2008; Taylor and Rohrer, 2010). On the other hand, if knowledge of whole anatomy functions as a precursor to building knowledge of sectional anatomy, as knowledge of arithmetic is a precursor to learning algebra, then the basic transfer (i.e., sequential) method should be best.

The present study investigated the possibility that an interleaved, trial by trial, transfer and integration between learning of whole and sectional anatomy (called the *alternation* condition), would lead to greater benefits for sectional anatomy learning than a basic transfer method (called the *whole then sections* condition). The efficiency of learning, generalization of knowledge to interpreting new sectional representations (Visible Human and MRI images), and long-term retention were measured.

This study had several differences from the earlier one. Chief among them were that a) the performance criterion for halting learning was relaxed from what had been a very demanding one, b) the retention interval for long-term retention was extended by an average of three weeks, and c) learning trials in the present study were blocked by anatomical view (e.g., coronal view; see later discussion). Due to these differences, it was decided to include a replication of the earlier study within this one. A third group was tested in which participants only learned sectional neuroanatomy. Comparison of the performance of that group (called *sections only*) with the whole then sections group provided an assessment of the degree of replication of the earlier study.

METHOD

Participants

Fifty nine undergraduate students at the University of Louisville were recruited for the study (22 male and 37 female). Participants were recruited through advertisements placed around the university campus and on a university website. Volunteers were surveyed about their knowledge of the neuroanatomical structures taught in the study. Only those with a rudimentary knowledge of neuroanatomy were enrolled. Participants were paid \$100 for completing the study.

Psychometric Test

Prior to the experiment, each participant was administered the Space Relations subtest of the Differential Aptitude Tests (Bennett et al., 1989). This is a standard test of spatial ability in which participants are required to indicate which of four solid figures (pictured in isometric views) would be formed by folding a flat design. Participants were assigned randomly to the three experimental conditions with the exception that the means and distributions of spatial ability scores were balanced across the learning groups.

MATERIALS

Participants learned individually using computer workstations with 24 inch high resolution color monitors. An accurate 3D computer graphical model of 19 structures in the human brain, and sections derived from the model, were developed for this project (Chariker et al., 2011). Images from the Visible Human, version 2.0, (Ackerman, 1995; Ratiu et al., 2003) made available by the National Library of Medicine were used as source information. These are photographs of cryosections through a cadaver brain. The structures that were modeled, and that could be selected in the learning programs, included amygdala, brainstem, caudate nucleus, cerebellum, cortex, fornix, globus pallidus, hypothalamus, hippocampus, mammillary body, nucleus accumbens, optic tract, pituitary, putamen, red nucleus, substantia nigra, subthalamic nucleus, thalamus, and ventricles.

One computer program supported learning whole brain anatomy (Figure 1A). In this case, the whole brain could be freely rotated and zoomed. Clicking on an anatomical structure with the computer mouse highlighted it and displayed its name. It then could be removed to reveal what was behind it (and later replaced).

A second program permitted exploration of sectional anatomy in any of the three standard anatomical views (i.e., coronal, sagittal, or axial; Figure 2A). Clicking on a structure highlighted it and displayed its name. Further exploration occurred through moving a slider that successively replaced the anatomical section that was displayed on the screen. (For more detail on program operation, see Chariker et al., 2011).

Both programs followed a sequence of free study for four minutes, a self-timed test, a numerical feedback panel, and feedback for two minutes. Testing in whole anatomy involved moving through the brain, selecting structures, and naming them by using a button panel (similar to that in Figures 1 and 2). Testing in sectional anatomy involved moving through sections of the brain, selecting structures that were indicated by red arrows, and naming them by using the button panel. Longer structures were tested in two positions (e.g., the head and tail of the caudate nucleus). Thus, there were 29 sectional test items for the 19 structures in the coronal and sagittal views, and 27 test items in the axial view. In both learning programs, items could be skipped, and items that had been named turned blue. There was no indication of which labels had been used, and labels could be used more than once.

The numerical feedback gave the number of items labeled correctly, the number mislabeled, the number omitted, and percentage correct. Graphical feedback permitted exploring the

brain exactly as in the study phase. However, correctly labeled structures were green, mislabeled structures were red, and omitted structures remained their standard colors (Figures 1B and 2B). Participants could click on any structure, including the mislabeled structures that were colored red, and immediately see the correct name.

Tests of long-term retention of sectional anatomy used a format identical to the learning trials, but without the study and feedback phases. Tests of generalization to the interpretation of biomedical images required labeling structures in MRI images and in Visible Human digital photographs. These tests also were conducted with interactive computer programs. For tests of long-term retention and generalization, every structure was tested in all three standard sectional views. As in sectional anatomy learning, there were 29 test items in the coronal and sagittal views, and 27 test items in the axial view.

For both sets of biomedical images (Visible Human and MRI), there were two different procedures for testing the generalization of knowledge to interpreting the images. In *Uncued Recognition*, participants clicked on any structure they recognized in the image displayed on the screen. A red dot was left on the screen where it was clicked. Then the appropriate name was selected in the button panel, and the pairing was submitted as an answer. The participants labeled as many structures as they could recognize in each image and then went to the next image. Sets of MRI and Visible Human images were selected so that all of the intended target test items would be available to the participant during the test. Responses were scored by comparing the locations of the red dots, and the labels, with images that had the target structures outlined and correctly labeled. Scorers were blind to the participants' experimental condition.

In the *Submit Name* test of generalization, a single structure in an image (Visible Human or MRI) was indicated by a red arrow, and the participant used the button panel to label the structure. Participants labeled one structure in each image and then proceeded to the next image. Because the data were labels associated with single images, they could be scored by computer. No feedback was given in the tests of generalization. For each type of test, all images of a single type (Visible Human or MRI) were tested together in a series.

Design and Procedure

As noted earlier, there were three experimental conditions. In *sections only*, participants learned only sectional neuroanatomy. Learning began with the coronal view and continued with that view until a test score reached at least 90% correct (17 of 19 structures). Learning trials then continued with the sagittal view until the 90% criterion was reached, and this was followed by trials with the axial view. After this sequence was completed, participants repeated it in a second cycle. Thus, the minimum number of trials to reach criterion was 6.

In the *whole then sections* condition, participants first learned whole anatomy according to the scheme just described: There were learning trials with each of the three anatomical views until test scores reached at least 90% correct in that view. The procedure began with the coronal view, progressed to the sagittal view, and ended with the axial view. This cycle was then repeated. For whole anatomy learning, the view determined the initial orientation of the brain model, but the model could be rotated. During the test phase of a whole anatomy

trial, the brain could only be rotated 45 degrees in any direction from the initial orientation (i.e., front, side, or top view).

After whole anatomy was learned in the whole then sections condition, sectional anatomy was learned with the same procedure (two cycles of the three anatomical views). Thus, the minimum number of trials for the whole then sections condition was 12: at least 90% correct in three anatomical views repeated over two cycles; once for whole anatomy and once for sectional anatomy.

The *alternation* condition followed a procedure similar to that of the whole then sections condition. However, whole and sectional anatomy were interleaved within single anatomical views (coronal, sagittal, or axial). The first trial within an anatomical view (e.g., the coronal view) presented whole anatomy. This was followed by a trial with sectional anatomy. The two types of presentation alternated until at least 90% correct was reached for whole anatomy, at which point presentation of whole anatomy ceased for that view. Trials with sectional anatomy continued for that view until at least 90% correct was reached. When performance reached at least 90% correct for both whole and sectional anatomy, learning moved to the next anatomical view (e.g., sagittal). When all three views had been learned in this fashion, learning trials proceeded through a second cycle of the three views. The minimum number of trials for the alternation condition again was 12: two cycles of trials for the three anatomical views in which at least 90% correct was reached for both whole and sectional anatomy.

Participants were given a thorough demonstration of the operation of the relevant instructional programs with a fictitious brain model. The learning trials did not begin until the individual could demonstrate and explain all elements of moving through a trial. Experimental sessions were limited to one hour, and participants were asked to schedule visits three times a week, as though they were attending a class. The two tests of generalization took place immediately after learning was completed. To minimize contamination between the tests, Uncued Recognition was always tested before Submit Name. Tests of long-term retention were conducted four to eight weeks later, depending on the participant's availability.

Statistical analyses were completed with the SPSS statistical package, version 20, (IBM Corp., Armonk, NY). All data sets were checked for extreme scores using the standard SPSS Explore program.

RESULTS

Mean test performance over learning trials is illustrated in Figure 3 for all three experimental conditions. Extreme scores from one participant were excluded from statistical analyses of the number of trials to complete learning.

Comparison of the Whole then Sections and Alternation Conditions

The primary concern in the present study was to compare learning in the WtS condition with learning in the alternation (Alt) condition. Two measures were of interest in assessing the

efficiency of learning for the two experimental conditions: the percent correct test performance in Trial 1 of each sectional view, and the number of trials required to complete each sectional view. As learning progressed from view to view, there should have been at least some transfer of learning for both experimental conditions. Thus, percent correct in Trial 1 of each view should have gone up, and the number of trials to complete learning in each view should have gone down. The mean percent correct in Trial 1 of each view, and the total number of trials to complete each view, are shown in Figure 4, broken down by experimental condition and anatomical view.

For the comparison of WtS and Alt, it was predicted that there would be an interaction of condition by view for these measures of learning. WtS would be superior to Alt at the first view, with higher percent correct in Trial 1 and fewer trials to complete learning. At the last anatomical view, Alternation would be superior on these measures, with higher percent correct in Trial 1 and fewer trials to complete learning.

Learning: Mean percent correct test performance, Trial 1 of each view—When averaged over both experimental conditions, test performance in Trial 1 of each anatomical view improved as learning progressed (Figure 4A), $F(2, 72) = 14.542, p < 0.001, \eta_p^2 = 0.288$. In pairwise comparisons, the first, coronal, view had lower test performance than each of the succeeding views: coronal vs. sagittal, $F(2, 38) = 77.502, p < 0.001, d = 1.65$; coronal vs. axial, $F(2, 38) = 73.905, p < 0.001, d = 1.70$.

The predicted statistical interaction of condition and view did occur for performance in Trial 1 of each view (Figure 4A), $F(2, 72) = 22.667, p < 0.001, \eta_p^2 = .386$. In the two planned comparisons, WtS had higher test performance than Alt in Trial 1 of the first (coronal) view, $t(37) = 4.705, p < .001, d = 1.51$; WtS $M = 67.6\%$, Alt $M = 45.7\%$. Thus, the prior learning of whole anatomy gave WtS a large advantage in the first trial of sectional anatomy learning. In the third and last (axial) view, the two groups were statistically equivalent.

Learning: Number of trials to complete each view—When averaged over both experimental conditions, the number of trials required to complete learning of each anatomical view decreased over the three views (Figure 4B), $F(2, 70) = 6.924, p = 0.002, \eta_p^2 = 0.165$. In pairwise comparisons, the first, coronal, view required significantly more trials than each of the succeeding views: coronal vs. sagittal, $F(2, 37) = 74.178, p < 0.001, d = 0.79$, coronal vs. axial, $F(2, 37) = 70.487, p < 0.001, d = 1.0$.

The predicted statistical interaction of condition and view occurred for the number of trials to complete each view (Figure 4B), $F(2, 70) = 4.095, p = 0.021, \eta_p^2 = 0.105$. In the two planned comparisons, WtS and Alt were statistically equivalent in the first view learned, $p = 0.345$. At the third and last view, Alt completed learning in fewer trials than WtS, $t(30.438) = 2.051, p = 0.049, d = 0.65$; WtS $M = 1.62$ trials, Alt $M = 2.34$ trials. Indeed, a striking outcome of this study was that more than half the participants in the alternation condition (55.6%) completed learning of the axial view of sectional anatomy in a single trial. The proportion of participants completing a view of sectional anatomy in a single trial is shown in Figure 5 for all three experimental groups.

Learning: number of total trials and errors to complete learning—The participants in the Alt condition completed all of neuroanatomy learning (whole and sectional anatomy) in 2 fewer trials than did WtS. Although the effect size for this comparison was moderate, this difference did not reach statistical significance, $t(36) = 1.447$, $P = 0.157$, $d = 0.47$; WtS $M = 20.1$ trials, Alt $M = 18.0$ trials.

Participants in the Alt condition committed 33% fewer errors (omissions and mislabeling summed over all tests completed) over the total course of neuroanatomy learning that did those in WtS, $t(28.72) = 2.051$, $P = 0.049$, $d = 0.65$; WtS $M = 88$ errors, Alternate $M = 59$ errors. Thus, although the test performance in Trial 1 of sectional anatomy learning was higher for WtS, Alt ultimately reached criterion across all three anatomical views with significantly fewer test errors.

Generalization to biomedical images—There were no statistically significant effects related to the difference between WtS and Alt for the tests of generalization to interpreting biomedical images. In regard to overall levels of performance, test scores for the Uncued Recognition test were relatively low, WtS $M = 34.3\%$, Alt $M = 33.8\%$. Performance was much better for the Submit Name test, WtS $M = 63.9\%$, Alt $M = 64.6\%$. The improvement can be explained by the fact that the biomedical images contained innumerable features that had not been present in learning (e.g., bone, muscle, blood vessels, and additional neural structures). When a structure was indicated to be part of the set that had been learned, performance in recognition improved substantially.

For recognition in the Submit Name test, participants were much better at recognizing structures in the higher resolution Visible Human cryosections ($M = 72.5\%$) than in the MRI images ($M = 56.0\%$); $F(1, 37) = 22.779$, $p < 0.001$, $\eta_p^2 = 0.381$. Thus, the resolution and contrast of the images also was a determinant of recognition performance in the tests of generalization.

Long term retention—The overall test performance for the long-term retention of sectional anatomy was 76.1%. There were no statistically significant effects related to the difference between WtS and Alt for the tests of long-term retention (WtS $M = 74.9\%$, Alt $M = 77.4\%$).

Replication: Comparison of Whole then Sections to Sections Only

The sections only condition was included in this study to assess the degree of replication of the earlier study, in which sections only (SO) and whole then sections (WtS) were compared (Chariker et al., 2011, 2012). We do not present the complete set of summaries and analyses in this report that would essentially repeat the earlier paper. Rather, we focus on important differences between the results of the two studies.

The comparison of WtS and SO did not replicate the earlier study in two places. The first instance is that the SO group in the present study required many fewer trials to complete learning than did the SO group in the earlier study (12.75 trials in the present study vs. 21.4 trials earlier). Moreover, the analogous reduction in trials was less for the present WtS group, because it had to meet the learning criterion for both whole and sectional anatomy

(Figure 3). Thus, when considering the number of total trials to complete all of neuroanatomy learning (whole and/or sectional anatomy), the present SO group required substantially fewer trials than did the WtS group: SO $M = 12.75$ trials, WtS $M = 20.05$ trials; $t(38) = 5.166$, $P < 0.001$, $d = 1.63$. In this study, WtS required 55% more trials than SO to complete all of learning. In the earlier study, WtS required 25% more trials than SO.

The second difference between the two studies was in the long term retention of sectional anatomy. Following the method of Chariker et al. (2012), the test items for retention of sectional anatomy were first divided statistically into those of typical difficulty and those that were especially difficult. The 85 total test items divided into 67 typical test items (78.8% of the items) and 18 difficult test items (21.2%). Each participant then received a score for retention of typical items and a score for retention of difficult items. There was a statistical interaction of experimental condition with item difficulty for long term retention of sectional anatomy, $F(1, 37) = 11.661$, $p = 0.002$, $\eta_p^2 = 0.240$. The WtS group had much higher retention of difficult structures than did the SO group, $F(1, 60.199) = 14.707$, $p < 0.001$, $d = .95$; WtS typical structures $M = 85.4\%$, SO typical structures $M = 83.9\%$, WtS difficult structures $M = 49.2\%$, SO difficult structures $M = 32.2\%$. This is a much larger advantage for WtS on long term retention of sectional anatomy than was seen earlier (Chariker et al., 2012).

DISCUSSION

This study was part of an effort to develop and test more efficient methods for instruction of neuroanatomy. Computer-based methods were developed on the belief that they present graphical material effectively and support the extensive self-study required in this discipline. Sectional neuroanatomy has been the primary target for instruction in this project due to its high level of difficulty. The computer programs have been tested in the context of a method called adaptive exploration, which takes advantage of the benefits for long-term retention of spaced practice, feedback, and repeated testing (Mahmood and Darzi, 2004; Larsen et al., 2008). The fundamental hypothesis behind the recent work has been that combining instruction of whole neuroanatomy with sectional neuroanatomy will be an effective way to provide organization and integration of the complex material that is presented in sectional neuroanatomy (see also Ruisoto et al., 2012).

The present study demonstrated that an integrative process of learning, in which whole and sectional neuroanatomy are presented in alternating trials, is a more efficient method of learning than the basic transfer paradigm. The typical participant in the alternation condition required only six cumulative trials (each with six minutes of study and feedback) to reach levels of test performance that remained above 80% correct for the three standard anatomical views of whole and sectional neuroanatomy. Participants in the alternation condition committed 33% fewer errors during neuroanatomy learning than did participants in the whole then sections condition. By the time participants in the alternation condition moved to the third (axial) view of sectional anatomy, more than half the participants were at or above 90% correct test performance in a single learning trial.

The improvement in the efficiency of learning for the interleaving of whole and sectional anatomy came at no cost to long-term retention or generalization to interpreting new biomedical images. We expect that as the amount of material to be learned is scaled up, and learning is extended over longer periods, the alternation method will increase in its relative effectiveness. There may then be benefits of alternation for long-term retention and generalization.

This study repeated a comparison from an earlier study (Chariker et al., 2011, 2012), in which a group that learned whole anatomy and then learned sectional anatomy (WtS) was compared with a group that learned only sectional anatomy (SO). The differences between the two studies are consistent with what is known about the relations between repeated testing and long term retention (Karpicke and Roediger, 2008; Larsen et al., 2008; Rohrer and Pashler, 2010). Participants in the SO group in the present study required relatively few trials to complete learning of sectional neuroanatomy. As part of this reduction, they took fewer tests. In conjunction with this outcome, this group had a greater disadvantage in long term retention of sectional neuroanatomy than was seen earlier. Thus, there appears to be a tradeoff. Framed in terms of the advantages of WtS over SO, WtS either will be: (1) relatively efficient in the number of trials required to learn both representations of neuroanatomy, with small benefits for long term retention of sectional anatomy, or (2) mastering two representations of neuroanatomy may require substantially more total learning trials than learning one, but long term retention of sectional anatomy will be superior (for further discussion of testing and retention in relation to this work, see Chariker et al., 2011).

There are several limitations of the present study. With only 20 participants in each group, statistical power was relatively low. A second limitation is that the instructional methods were not compared to alternative educational technologies, such as the use of textbooks and atlases. Our approach to this problem has been that new technologies need to advance in their development before it is useful to compare them to alternatives (consider also Friedman, 1994; Cook, 2005). Meanwhile, tests of external validity, such as the recognition of structures in biomedical images, can be used as benchmarks to monitor the progress of development. Finally, only 19 neuroanatomical structures were taught. Actual instruction of neuroanatomy includes many more structures, several more subdomains of anatomy, and neural function. Thus, this study serves as an initial trial for developing methods of instruction in computer-based systems. Scaling the methods up is an important challenge for future work.

ACKNOWLEDGEMENTS

Primary support for this research came from grant R01 LM008323 from the National Library of Medicine, National Institutes of Health, (PI: J. Pani). The research also was supported by a grant from the National Science Foundation and Defense Intelligence Agency, USA (IIS-0650138). The authors thank Cassandra Marshall, Mayur Patel, and Katherine Schmitt for help with data collection. We thank the Surgical Planning Laboratory in the Department of Radiology at Brigham and Women's Hospital and Harvard Medical School for use of MRI images from the SPL-PNL Brain Atlas. Preparation of the atlas was supported by NIH grants P41 RR13218 and R01 MH050740. We thank the National Library of Medicine for use of the Visible Human 2.0 photographs.

Biographies

JOHN R. PANI, Ph.D., is a professor in the Department of Psychological and Brain Sciences and the Bioinformatics Core at the University of Louisville, Louisville, Kentucky. He teaches courses in human cognition, cognitive neuroscience, learning, and history of psychology. His research combines the psychology of learning and cognition, new information technologies, and collaboration with experts in biology and medicine.

JULIA H. CHARIKER, Ph.D., is a postdoctoral scholar in the Department of Psychological and Brain Sciences and the Bioinformatics Core at the University of Louisville, Louisville, Kentucky. She teaches courses in human cognition and learning. Her research combines the psychology of learning and cognition, new information technologies, and collaboration with experts in biology and medicine.

FARAH NAAZ, Ph.D., was a graduate student in the Department of Psychological and Brain Sciences at the University of Louisville, Louisville, Kentucky, during conduct of this research. She is now a postdoctoral research associate in the Department of Psychology at Washington University in St. Louis, St. Louis, Missouri. Her research focuses on visual cognition, learning, and memory.

LITERATURE CITED

- Ackerman, MJ. D-Lib Magazine. Reston, VA: Corporation for National Research Initiatives; 1995. Accessing the Visible Human project. URL: <http://www.dlib.org/dlib/october95/10ackerman.html> [accessed 8 February 2012]
- Adams CM, Wilson TD. Virtual cerebral ventricular system: An MR-based three-dimensional computer model. *Anat Sci Educ.* 2011; 4:340–347. [PubMed: 21976457]
- Bennett, GK.; Seashore, HG.; Wesman, AG. *Differential Aptitude Tests[®] for Personnel and Career Assessment: Space Relations.* San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovich; 1989.
- Bower GH. Organizational factors in memory. *Cognit Psychol.* 1970; 1:18–46.
- Bower GH, Clark MC, Lesgold AM, Winzenc D. Hierarchical retrieval schemes in recall of categorized word lists. *J Verbal Learn Verbal Behav.* 1969; 8:323–343.
- Bransford, JD.; Brown, AL.; Cocking, RR., editors. *How People Learn: Brain, Mind, Experience, and School.* 2nd Ed.. Washington, DC: National Academy Press; 2000. p. 385
- Cepeda NJ, Pashler H, Vul E, Wixted JT, Rohrer D. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychol Bull.* 2006; 132:354–380. [PubMed: 16719566]
- Chariker JH, Naaz F, Pani JR. Computer-based learning of neuroanatomy: A longitudinal study of learning, transfer, and retention. *J Educ Psychol.* 2011; 103:19–31. [PubMed: 23349552]
- Chariker JH, Naaz F, Pani JR. Item difficulty in the evaluation of computer-based instruction: An example from neuroanatomy. *Anat Sci Educ.* 2012; 5:63–75. [PubMed: 22231801]
- Collins JP. Modern approaches to teaching and learning anatomy. *BMJ.* 2008; 337:665–667.
- Cook DA. The research we still are not doing: An agenda for the study of computer-based learning. *Acad Med.* 2005; 80:541–548. [PubMed: 15917356]
- Crowley RS, Medvedeva O. An intelligent tutoring system for visual classification problem solving. *Artif Intell Med.* 2006; 36:85–117. [PubMed: 16098717]
- Crowley RS, Naus GJ, Stewart J 3rd, Friedman CP. Development of visual diagnostic expertise in pathology: An information processing study. *J Am Med Inform Assoc.* 2003; 10:39–51. [PubMed: 12509356]
- Friedman CP. The research we should be doing. *Acad Med.* 1994; 69:455–457. [PubMed: 8003158]

- Höhne, KH.; Petersik, A.; Pflesser, B.; Pommert, A.; Priesmeyer, K.; Riemer, M.; Schiemann, T.; Schubert, R.; Tiede, U.; Urban, M.; Frederking, H.; Lowndes, M.; Morris, J. *Regional, Functional, and Radiological Anatomy (DVD)*. New York NY: Springer Electronic Media; 2003. *Voxel-Man 3D-Navigator: Brain and Skull*.
- Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. *Med Teach*. 2005; 27:10–28. [PubMed: 16147767]
- Karpicke JD, Roediger HL 3rd. The critical importance of retrieval for learning. *Science*. 2008; 319:966–968. [PubMed: 18276894]
- Kockro RA, Stadie A, Schwandt E, Reisch R, Charalampaki C, Ng I, Yeo TT, Hwang P, Serra L, Perneckzy A. A collaborative virtual reality environment for neurosurgical planning and training. *Neurosurgery*. 2007; 61:379–391. [PubMed: 18091253]
- Kornell N, Bjork RA. Learning concepts and categories: Is spacing the “enemy of induction”? *Psychol Sci*. 2008; 19:585–592. [PubMed: 18578849]
- Kornell N, Hays MJ, Bjork RA. Unsuccessful retrieval attempts enhance subsequent learning. *J Exp Psychol Learn Mem Cogn*. 2009; 35:989–998. [PubMed: 19586265]
- Larsen DP, Butler AC, Roediger HL 3rd. Test-enhanced learning in medical education. *Med Educ*. 2008; 42:959–966. [PubMed: 18823514]
- Lesgold, A.; Rubinson, H.; Feltovich, P.; Glaser, R.; Klopfer, D.; Wang, Y. Expertise in a complex skill: Diagnosing x-ray pictures. In: Chi, MTH.; Glaser, R.; Farr, MJ., editors. *The Nature of Expertise*. 1st Ed.. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1988. p. 311-342.
- Luursema J-M, Verwey WB, Kommers PA, Geelkerken RH, Vos HJ. Optimizing conditions for computer-assisted anatomical learning. *Interact Comput*. 2006; 18:1123–1138.
- Mahmood T, Darzi A. The learning curve for a colonoscopy simulator in the absence of any feedback. No feedback, no learning. *Surg Endosc*. 2004; 18:1224–1230. [PubMed: 15457382]
- Malone HR, Syed ON, Downes MS, D'Ambrosio AL, Quest DO, Kaiser MG. Simulation in neurosurgery: A review of computer-based simulation environments and their surgical applications. *Neurosurgery*. 2010; 67:1105–1116. [PubMed: 20881575]
- Nicholson DT, Chalk C, Funnell WR, 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. [PubMed: 17054617]
- Nowinski WL, Thirunavuukarasuu A, Ananthasubramaniam A, Chua BC, Qian G, Nowinska NG, Marchenko Y, Volkau I. Automatic testing and assessment of neuroanatomy using a digital brain atlas: Method and development of computer- and mobile-based applications. *Anat Sci Educ*. 2009; 2:244–252. [PubMed: 19743409]
- Pani JR, Chariker JH, Dawson TE, Johnson N. Acquiring new spatial intuitions: Learning to reason about rotations. *Cogn Psychol*. 2005a; 51:285–333. [PubMed: 16165120]
- Pani, JR.; Chariker, JH.; Fell, RD. Visual cognition in microscopy. In: Bara, BG.; Barsalou, L.; Bucciarelli, M., editors. *Proceedings of the XXVII Annual Conference of the Cognitive Science Society (CogSci 2005)*. 1st Ed.. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.; 2005b. p. 1702-1707.
- Pashler H, Cepeda NJ, Wixted JT, Rohrer D. When does feedback facilitate learning of words? *J Exp Psychol Learn Mem Cogn*. 2005; 31:3–8. [PubMed: 15641900]
- Petersson H, Sinkvist D, Wang C, Smedby O. Web-based interactive 3D visualization as a tool for improved anatomy learning. *Anat Sci Educ*. 2009; 2:61–68. [PubMed: 19363804]
- Primal Pictures. *3D Human Anatomy Software*. London, UK: Primal Pictures Ltd; 2007. URL: www.primalpictures.com [accessed 2 October 2011]
- Ratiu P, Hillen B, Glaser J, Jenkins DP. Visible Human 2.0--the next generation. *Stud Health Technol Inform*. 2003; 94:275–281. [PubMed: 15455907]
- Rohrer D, Pashler H. Recent research on human learning challenges conventional instructional strategies. *Educ Res*. 2010; 39:406–412.
- Rohrer D, Taylor K. The shuffling of mathematics problems improves learning. *Instr Sci*. 2007; 35:481–498.

- Ruisoto P, Juanes JA, Contador I, Mayoral P, Prats-Galino A. Experimental evidence for improved neuroimaging interpretation using three-dimensional graphic models. *Anat Sci Educ.* 2012; 5:132–137. [PubMed: 22434672]
- Ruiz JG, Cook DA, Levinson AJ. Computer animations in medical education: A critical literature review. *Med Educ.* 2009; 43:838–846. [PubMed: 19709008]
- Solomon B, Bizekis C, Dellis SL, Donington JS, Olikier A, Balsam LB, Zervos M, Galloway AC, Pass H, Grossi EA. Simulating video-assisted thoracoscopic lobectomy: A virtual reality cognitive task simulation. *J Thorac Cardiovasc Surg.* 2011; 141:249–255. [PubMed: 21168026]
- Sugand K, Abrahams P, Khurana A. The anatomy of anatomy: A review for its modernization. *Anat Sci Educ.* 2010; 3:83–93. [PubMed: 20205265]
- Tam MD, Hart AR, Williams S, Heylings D, Leinster S. Is learning anatomy facilitated by computer-aided learning? A review of the literature. *Med Teach.* 2009; 31:e393–e396. [PubMed: 19811174]
- Taylor K, Rohrer D. The effects of interleaved practice. *Appl Cognit Psychol.* 2010; 24:837–848.
- Trelease RB. Diffusion of innovations: Smartphones and wireless anatomy learning resources. *Anat Sci Educ.* 2008; 1:233–239. [PubMed: 19109851]
- UW-SIG. The digital anatomist information system. Seattle, WA: University of Washington; 1999. University of Washington, Structural Informatics Group, Department of Biological Structure, University of Washington. URL: <http://sig.biostr.washington.edu/projects/da/> [accessed 8 February 2012]
- Winkelman A. Anatomical dissection as a teaching method in medical school: A review of the evidence. *Med Educ.* 2007; 41:15–22. [PubMed: 17209888]
- Yeung JC, Fung K, Wilson TD. Development of a computer-assisted cranial nerve simulation from the visible human dataset. *Anat Sci Educ.* 2011; 4:92–97. [PubMed: 21438158]

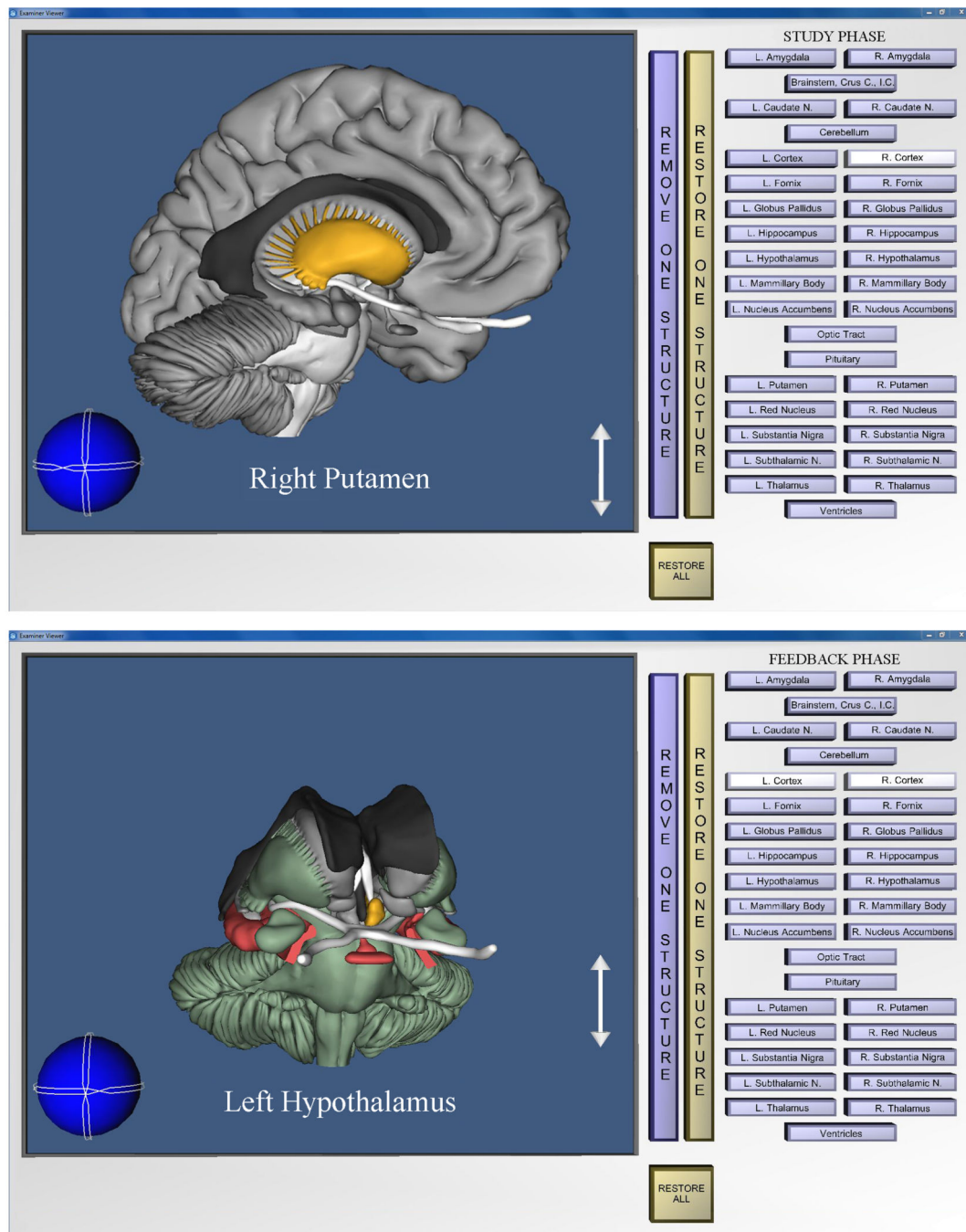


Figure 1. Screen images from the program for learning whole anatomy. **A:** Image from the study phase of a trial. The learner has rotated the brain, removed the right cortical hemisphere, and highlighted the right putamen. **B:** Image showing post-test graphical feedback. The learner has removed the cortical hemispheres and highlighted the left hypothalamus. Structures colored green were labeled correctly on the test, structures colored red were mislabeled, and achromatic structures were skipped.

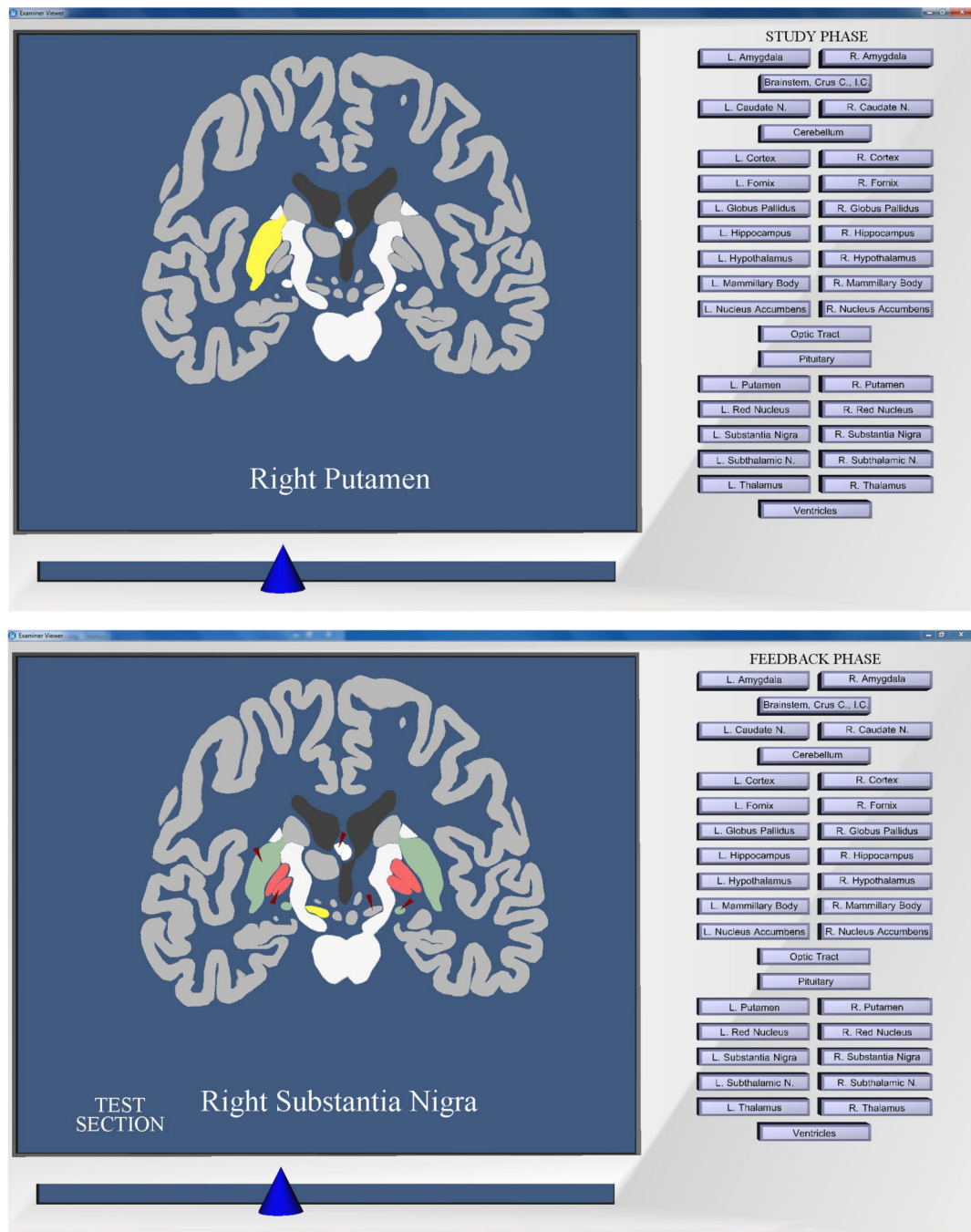


Figure 2.

Screen images from the program for learning sectional anatomy. **A:** Image from the study phase of a trial. The learner has moved to a coronal section that cuts through the middle of the brain and has highlighted the right putamen. **B:** Image showing post-test graphical feedback. Tested items are indicated by small red arrows. Structures colored green were labeled correctly on the test, structures colored red were mislabeled, and achromatic structures indicated by the red arrows were skipped. The learner has highlighted the right substantia nigra.

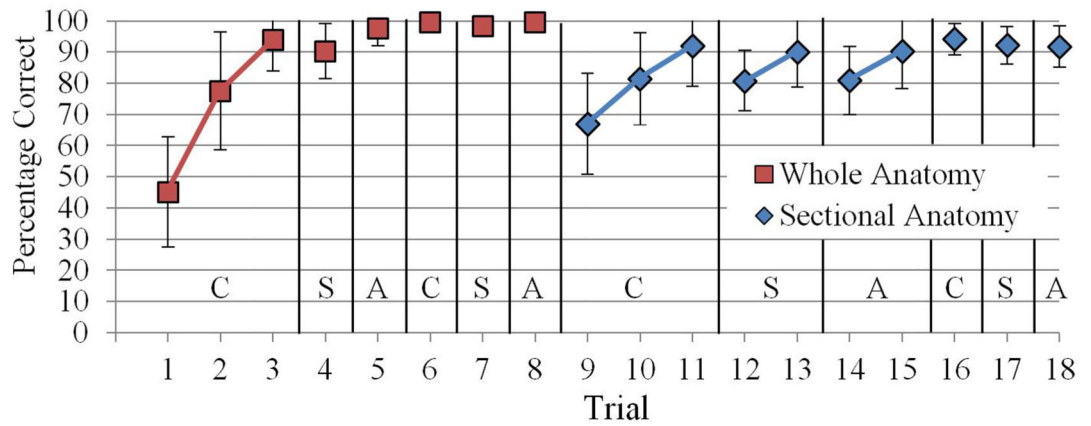
Author Manuscript

Author Manuscript

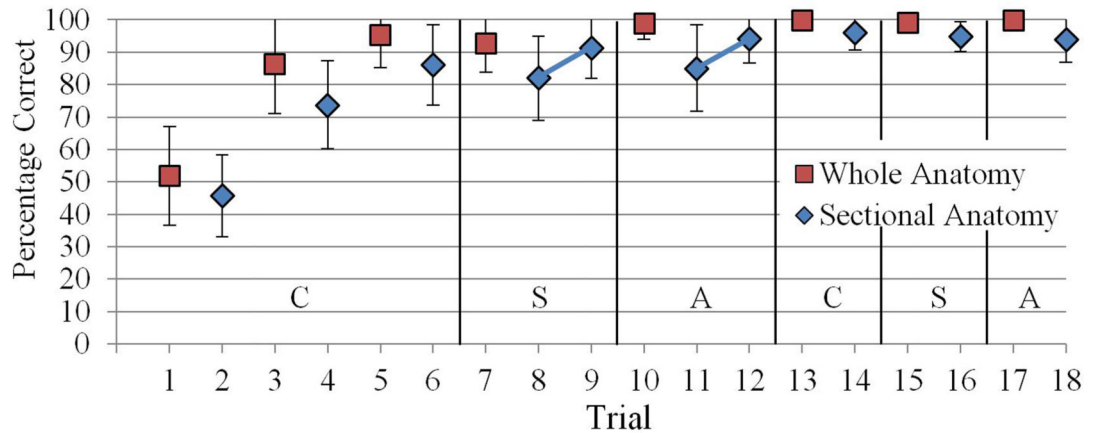
Author Manuscript

Author Manuscript

Whole then Sections



Alternation



Sections Only

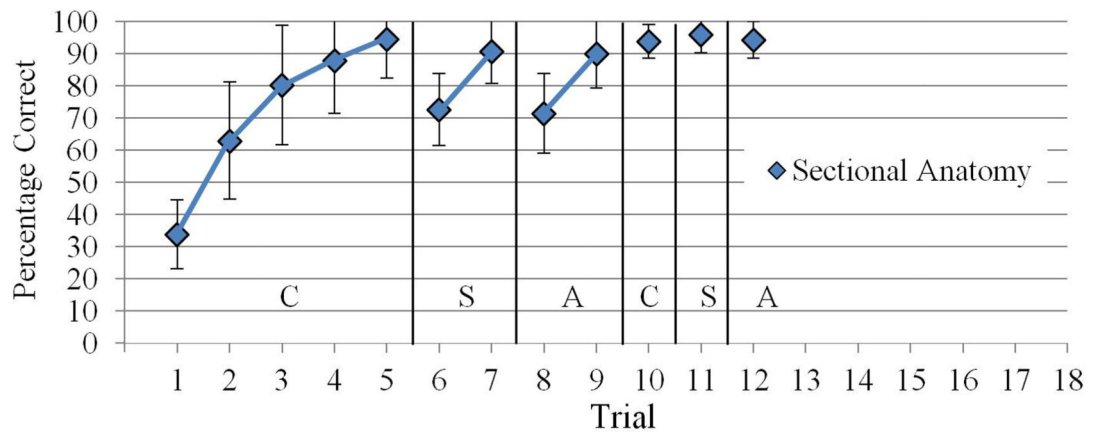


Figure 3.

Mean percentage correct test performance across learning trials, broken down by type of anatomical presentation and experimental condition. C = Coronal section. S = Sagittal section. A = axial section. Error bars illustrate sample standard deviations. Note that not all participants completed the same number of trials. Where means represent a trial in which some participants had met the performance criterion earlier, those participants' performance was extrapolated forward to contribute to the group mean.

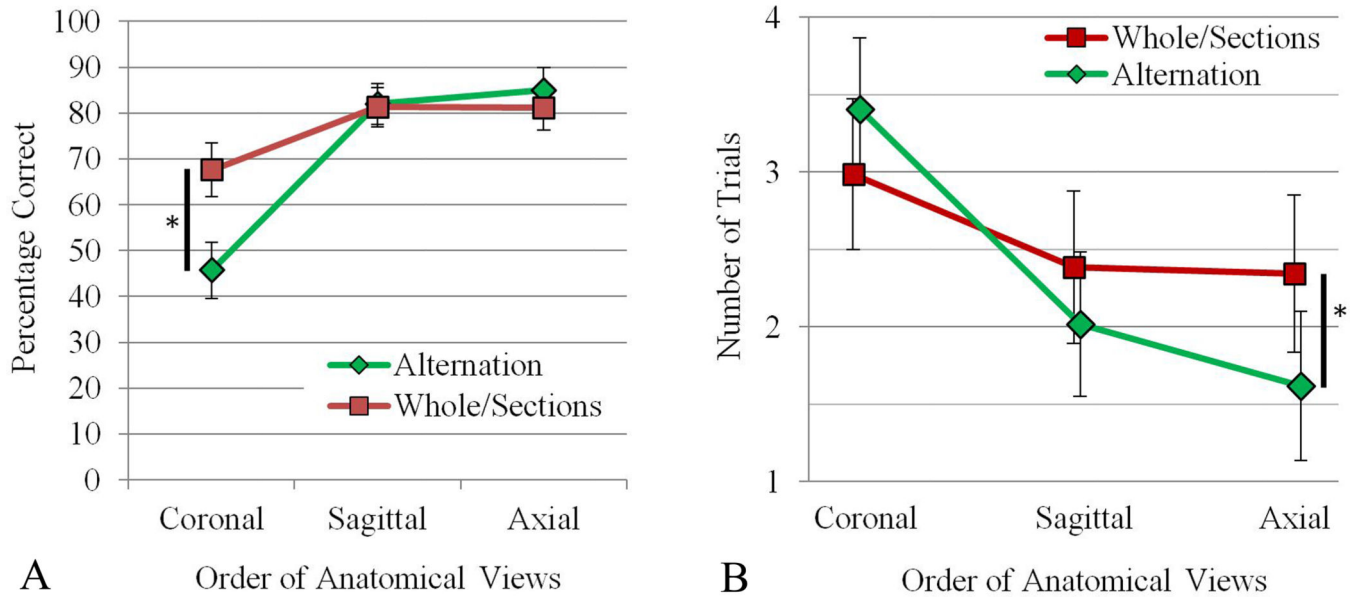


Figure 4.

A: Mean percentage correct test performance in the first trial of each anatomical view for the Alternation and Whole then Sections groups. **B:** The number of trials required to reach the performance criterion for each anatomical view for the Alternation and Whole then Sections groups. Vertical black bars with an asterisk indicate a statistically significant pairwise comparison ($P < 0.05$). Error bars show 95% confidence intervals.

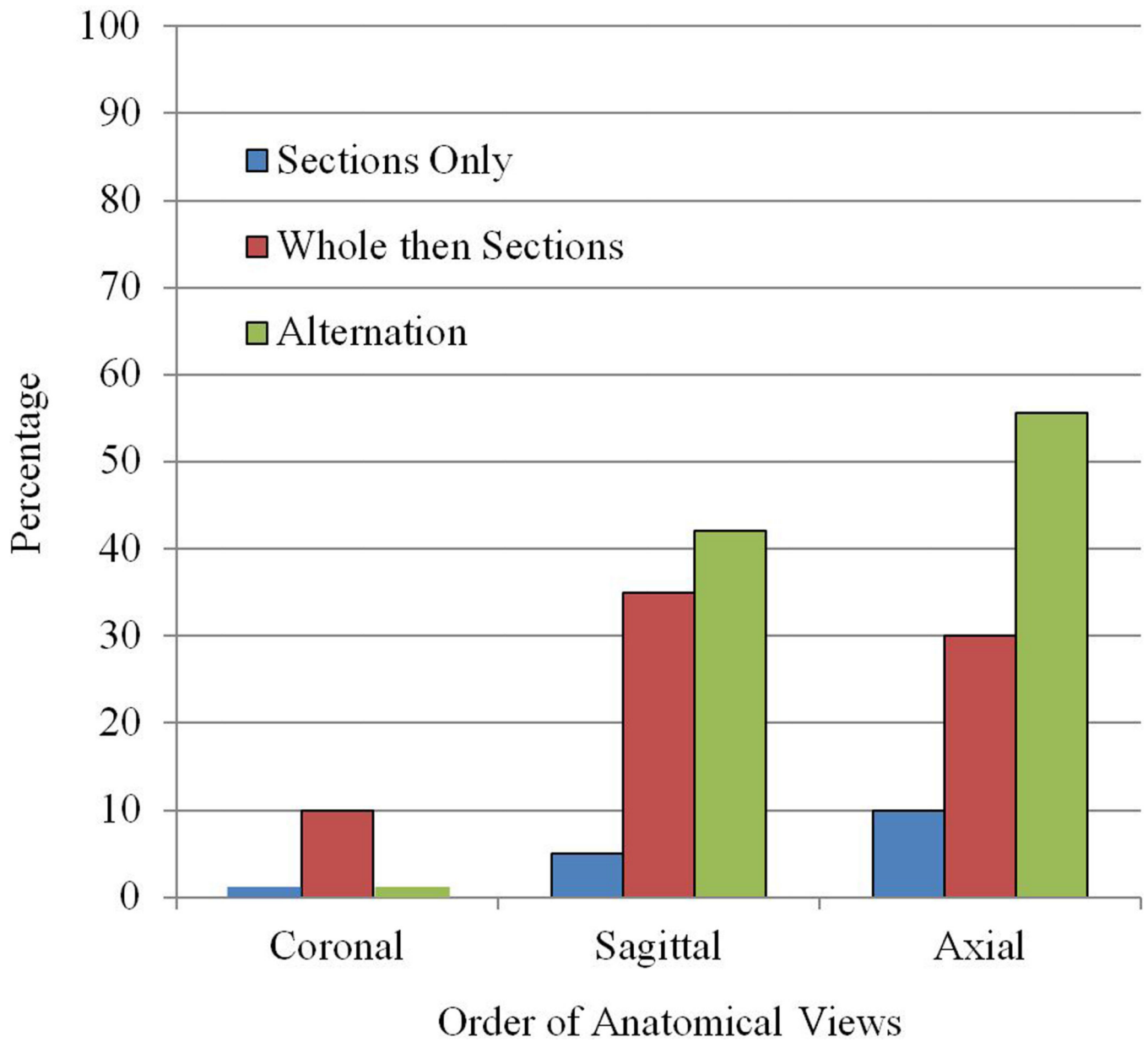


Figure 5. The percentage of participants in each of the three experimental groups who completed sectional anatomy learning for an anatomical view in one trial.