**Erolin, C, Lamb, C, Soames, R and Wilkinson, C**

# Does Virtual Haptic Dissection Improve Student Learning? A Multi-Year Comparative Study.

# Does Virtual Haptic Dissection Improve Student Learning? A Multi-Year Comparative Study

Caroline EROLIN [1, a], Clare LAMB [a], Roger SOAMES [a] and Caroline WILKINSON [b]
[a] *Centre for Anatomy and Human Identification, University of Dundee*
[b] *Face Lab, Liverpool John Moores University*

**Abstract.** This study investigated the haptic 'dissection' of a digital model of the hand and wrist in anatomy education at both undergraduate (UG) and postgraduate (PG) levels. The study ran over five successive years and was split into three discreet phases. Phase one compared the results of PG students across control, non-haptic and haptic groups. Phase two compared the results of UG students between control and haptic groups. Phase three compared the results of UG students across control, non-haptic and haptic groups. Results for all phases indicate that use of the model, both through haptic and non-haptic interfaces produced some significantly improved test results. The non-haptic group performing the strongest overall indicating that the addition of haptic feedback may not be beneficial to student learning.

**Keywords.** Anatomy, dissection, haptic, virtual, 3D visualization,

## 1. Introduction

With decreasing hours being dedicated to the teaching of anatomy in the UK and US [1, 2], as well as access to cadavers becoming more scarce [3, 1, 4], many universities are turning to digital resources. The past decade has seen the release of a number of computer packages aimed at enhancing anatomical education. However, there has been a paucity of research into the benefit or otherwise of these packages for student education.

In previous 'virtual anatomy' research, the primary emphasis has been on the replacement of cadaveric dissection [4, 5, 6], with only a few exceptions highlighting the benefits of integrating new technologies with existing teaching practice [3, 7, 8]. It is interesting to note, that although many existing software packages offer the replacement of cadaveric dissection through interactive three-dimensional anatomical models, none offers true virtual dissection, i.e. cutting through the anatomical layers with a haptic (tactile) interface.

The study of anatomy and dissection in particular is multisensory. It combines the act of 'doing' (cutting etc.) with a highly visual experience. A technology that allows the user to experience anatomy using the same multiple senses as in reality offers the possibility of being a very useful educational tool.

It was with this in mind that the current research aimed to study the potential benefit of integrating new technologies with existing methods of teaching. Consequently, a

[1] Corresponding Author. c.d.erolin@dundee.ac.uk

three-dimensional digital model of the hand and wrist was created which could be virtually ‘dissected’ through a haptic interface. The model was used as a teaching and revision aid both prior to and following cadaveric dissection. A haptic enabled version of the model, allowing for real-time dissection, was compared with a non-haptic version using a keyboard and mouse. Both versions were tested on students of gross anatomy and compared with respect to test results and student experience.

## 2. Methodology

### *2.1. Creating the Digital Model*

A digital model of the hand and wrist was created using Computerised Tomography (CT) and photographic data from the Visible Human Project (VHP) Female [9]. The 3D visualization software Amira® (5.2.2) was used to automatically separate out the hard tissues as well as manually segment the soft tissues. The hard tissues were automatically created from the CT data by altering the threshold value used to create an isosurface. The resulting 3D reconstruction was saved as an STL file.

Photographic cross-sections of the VHP female were used to create the soft tissues. However, the soft tissues of the hand and wrist are too small and similar, in terms of density and colour gradient, to enable the creation of isosurfaces for each structure in the same way as for the CT data. Each structure needed to be exported as a separate STL file to allow individual manipulation in the 3D modelling software Geomagic Freeform® (or ‘Sensable Freeform Modelling’ as it was known at the time). In order to achieve this each structure was manually ‘segmented’ in Amira® via the labelField window, with the structure being manually selected on every slice and added to the materials list. Once segmentation was complete the individual structures were saved and exported as STL files. Figure 1 shows the ‘surfaceview’ of the segmented structures in Amira.

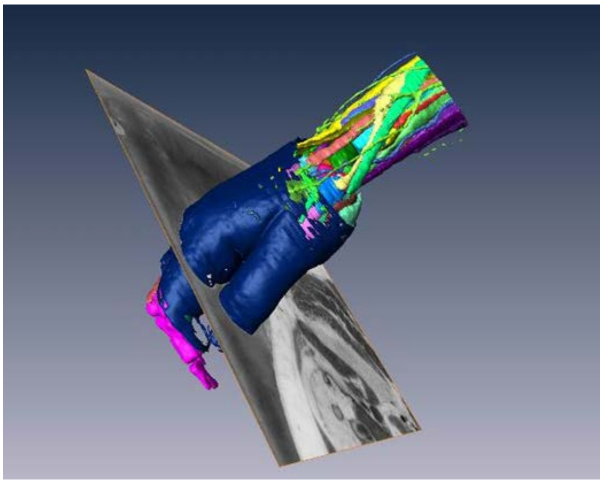


**Figure 1.** Screen shot from Amira® of ‘surfaceview’ for all structures.

Each structure was exported from Amira® with location coordinates so that they retained their spatial relationships when opened in Freeform®. When imported into

Freeform® each individual structure required varying degrees of remodelling where detail had been lost during the segmentation process. This involved smoothing out jagged edges as well as remodelling missing and partial elements.

The nerves, veins and arteries were difficult to observe and therefore segment on the cross-sectional images. Small elements were segmented and exported and acted as a guide for freehand modelling. Some details and fine structures could not be visualised as they are presented in the cadaver due to the limitations of the modelling system. These included very small branches of the nerves and vessels as well the fascia surrounding muscles and other structures.

The skin was created from the VHP CT data as an isosurface and exported as a solid STL file which was ‘shelled’ within Freeform®. The fat was created from a duplicate of the solid skin layer. It is possible within Freeform® to remove one area of virtual clay from another: the skin, muscles, nerves and vessels were therefore removed from the fat layer to leave only material in the voids between other structures.

The interface for interaction with the model was an adaptation of the Freeform® system used to create it. The software allows alterations to be made to the desktop layout so that all unnecessary tools and windows can be hidden, leaving only those required for interaction with the model.

Interaction with the model was possible through either a keyboard and mouse or haptic interface (Geomagic touch or touch x). Use of a haptic interface not only adds the sense of touch, but also alters the way in which the user interacts with and views the model. For example, cutting through the layers of the model produces a ‘window of view’ similar to that produced during traditional cadaveric dissection (Figure 2).

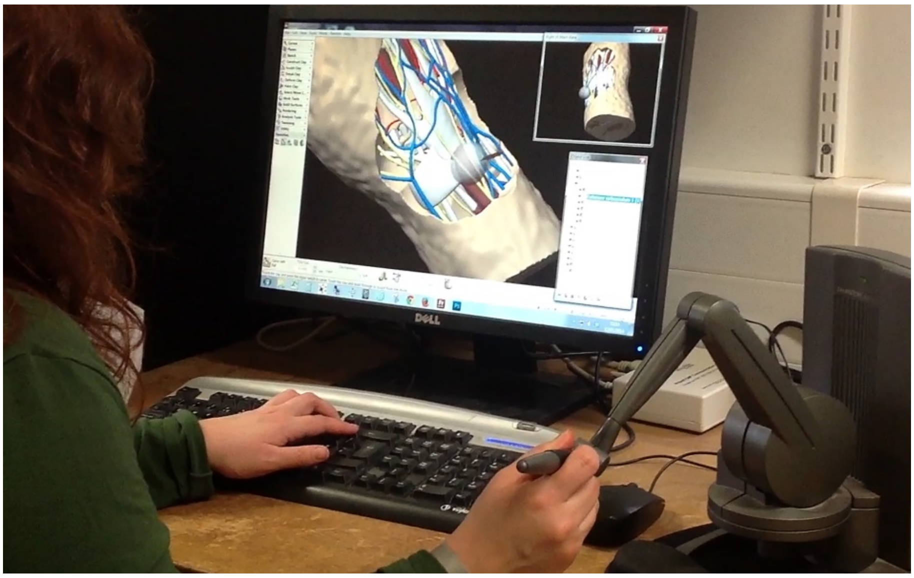



**Figure 2.** ‘Window of view’ created by cutting with the haptic device (Geomagic touch x).

### *2.2. Research Design*

The potential benefit of the software as an aid to learning anatomy and to improving cadaver dissection was assessed by testing both haptic-enabled and non-haptic versions of the software. Both versions were tested on anatomy students and compared with respect to test results (which did not contribute to their grades), dissection quality and student experience: only the test results are discussed here. The software was made available in addition to their normal tuition.

The study ran over five successive years and has been divided into three discreet phases. Phase one took place over four years (2011-2014) and compared the results of postgraduate MSc (PG) students across control (2011/12), non-haptic (2013) and haptic (2014) groups. Phase two took place over one year (2014) and compared undergraduate BSc (UG) students across control and haptic groups. Phase three took place over one year (2015) and compared the undergraduate BSc (UG) students across control, non-haptic and a haptic groups.

## 3. Results

### *3.1. Phase One Test Results*

The first set of data from PG anatomy students was collected between 2011 and 2014. The control group consisted of thirteen students while the non-haptic and haptic groups consisted of seven students each. The anatomy test consisted of eight questions, with questions one to three being anatomy 'spotters' i.e. identification questions, while questions four to eight were multiple choice. Statistical analysis was performed using the Mann-Whitney U test in SPSS (version 23).

Question one was an identification test consisting of nine elements: the non-haptic group scored highest with a total of 74.6%, followed by the control group with 63.2%, and finally the haptic group with 60.3%. None of these differences were significant. Question two was also an identification test consisting of six elements: the non-haptic group scored highest with a total of 80.9%, followed by the haptic group with 73.8%, and finally the control group with 52.5%. The difference between the control and non-haptic groups, and between the control and haptic groups were both statistically significant, with p-values of 0.003 and 0.030 respectively. Question three was another identification test consisting of ten elements: the non-haptic group scored highest with a total of 84.2%, followed by the haptic group with 71.4% and finally the control group with 41.6%. The difference between the control and non-haptic groups was again significant, with a p-value of 0.030. Questions 4 to 8 were multiple-choice. The non-haptic and control groups scored a total (for all four questions combined) of 80% each, followed by the haptic group with 77.1%. The difference was not significant. (Table 1).

### *3.2. Phase Two Test Results*

The second set of data was from UG anatomy students and was collected during 2014. The 4th (final) year students acted as a control group (they had dissected the hand and wrist the previous year and been given a week for revision prior to the test) with the 3rd year students having access to the haptic-enabled version of the software. The control and haptic groups consisted of seven and fourteen students respectively. The UG students were given the same anatomy test as the PG students. Statistical analysis was again performed using the Mann-Whitney U test in SPSS.

In question one the haptic group scored highest with a total of 52.3% followed by the control group with 39.7%: the difference was not significant. In question two the haptic group scored highest with a total of 54.7% followed by the control group with 19%: the difference was significant with a p-value of 0.004. In question three the haptic group scored highest with a total of 45.7% followed by the control group with 38.5%: the difference was not significant. For questions 4 to 8 the haptic group scored highest with a total of 82.8% followed by the control group with 62.8%: the difference was significant with a p-value of 0.010. (Table 1).

*3.3. Phase Three Test Results*

The final set of data was from UG anatomy students and was collected during 2015. The 4th (final) year students acted as a control group (they had dissected the hand and wrist the previous year and been given a week for revision prior to the test) with the 3rd year students being split into a further two groups; one having access to a non-haptic version of the software and one to a haptic-enabled version. The control consisted of twenty students, the non-haptic of five and haptic group of six. The UG students were given the same anatomy test as the PG students. Statistical analysis was again performed using the Mann-Whitney U test in SPSS.

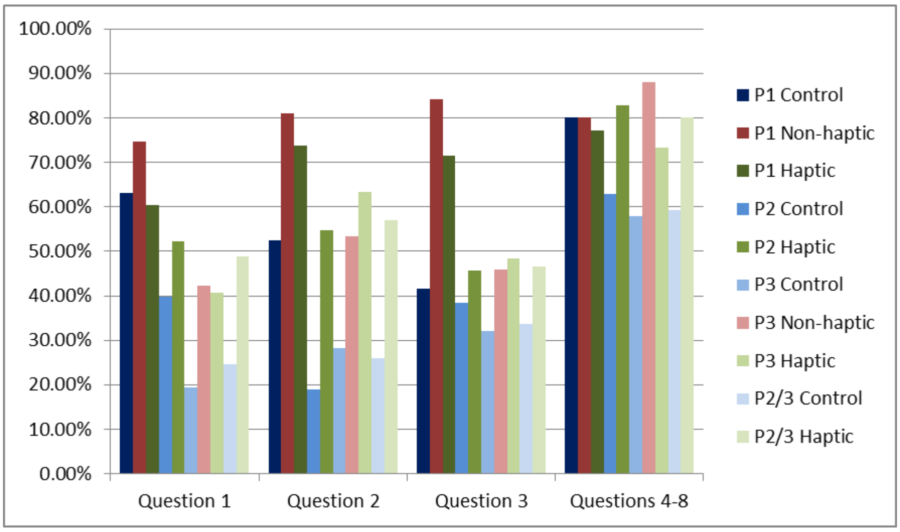
In question one the non-haptic group scored highest with a total of 42.2% followed by the haptic group with 40.7% and finally the control group with 19.4%: none of these differences were significant. In question two the haptic group scored highest with a total of 63.3% followed by the non-haptic group with 53.3% and finally the control group with 28.3%: none of these differences were significant. In question three the haptic group scored highest with a total of 48.3% followed by the non-haptic group with 46% and finally the control group with 32%: none of these differences were significant. For questions 4 to 8 the non-haptic group scored highest with a total of 88% followed by the haptic group with 73.3% and finally the control group with 58%: the difference between the control and the non-haptic group was significant with a p-value of 0.007, while the difference between the control and the haptic group was not significant. (Table 1).

*3.4. Phases Two and Three Combined*

The data from phases two and three was combined to create larger groups for further analysis. The combined control group (2014 and 2015) consisted of twenty seven students, the non-haptic group (2015 only) of five and the combined haptic group (2014 and 2015) of twenty students.

In question one the haptic group scored highest with a total of 48.8% followed by the non-haptic group with 42.2% and finally the control group with 24.7%: the difference between the control and the non-haptic group was not significant, however the difference between the control and the haptic group was significant with a p-value of 0.001. In question two the haptic group scored highest with a total of 57% followed by the non-haptic group with 53.3% and finally the control group with 25.9%: the difference between the control and the non-haptic group was not significant, however the difference between the control and the haptic group was significant with a p-value of 0.001. In question three the haptic group scored highest with a total of 46.5% followed by the non-haptic group with 46% and finally the control group with 33.7%: the difference between the control and the non-haptic group was not significant, however the difference between the control and the haptic group was significant with a p-value of 0.076. For questions 4 to 8 the non-haptic group scored highest with a total of 88% followed by the haptic group with 80% and finally the control group with 59.2%: The difference between the control and non-haptic group, and between the control and haptic group were both statistically significant, with p-values of 0.008 and 0.000 respectively. (Table 1).

**Table 1.** Results for all groups across all three phases of the study.




## 4. Discussion and Conclusion

### *4.1. Phase One (PG)*

The non-haptic group scored significantly higher than the control in two of the four question groups, while the haptic group scored significantly higher once. Some insight as to why this might be is given in the feedback comments where a number of participants suggested that while the haptic feedback device was interesting to use, it might actually be getting in the way of learning. For example one participant commented: "*I found the haptic device difficult to use, the tool kept getting stuck inside the hand and the resistance made me tire quickly and my wrist sore. After a short time with the device I resorted to using the mouse and keyboard*".

Cognitive Load Theory (CLT) likely goes some way to explaining this. Cognitive load is defined as the load that performing a task imposes on the learner's cognitive system. CLT suggests that the learning process takes up a large portion of working memory and that if too much information is presented simultaneously the working memory may be overloaded, impeding learning [10]. What the educator brings to the student is considered either extraneous or germane load. Extraneous load is usually the result of badly delivered or unnecessary information, resulting in the student having to use additional thought processes to identify the relevant material [10]. Learning to use an additional piece of novel hardware, such as a haptic device, in order to learn anatomy from a 3D digital model might be considered extraneous load. Despite this there were also a number of positive comments regarding the haptic interface, for example: "*It was odd to get used to the resistance and* *being able to feel the surfaces, but it helped to make (me) realise that it was 3D and the thumb structures would be more anteriorly placed"*.

### *4.2. Phase Two (UG)*

The haptic group scored significantly higher than the control in two of the four question groups. It was suggested in the discussion of phase one that the haptic device could potentially be 'getting in the way' of student learning by overloading their working memory (as per CLT). Whether this was the case here is difficult to say with no non-haptic group for comparison. However, feedback does suggest that this may have been the case for some students at least, for example: "*I did try to use the haptic interface, but I feel I would need much more time with it in order to be quite confident using it"*.

However, there were also a large number of comments suggesting that for some the addition of the haptic device was helpful and engaging: "*It was good to physically feel the difference and where you were 'cutting"'. "The haptic interface was very useful as you could literally feel as you were dissecting. This aided real life dissection". "It was good to be able to 'feel' the cadaver and it probably helped my dissection skills"*. From these comments, it appears that some students found the haptic device more intuitive and easy to use than others. This could be related to the individual's spatial ability, as well as prior exposure to 3D digital models. This information was not recorded in this study but could be valuable in future research.

Previous studies [11, 12] have found that anatomical knowledge involves spatial reasoning in three dimensions and that students who perform poorest in spatial exercises tend to score significantly lower in practical anatomy examinations. Studies have shown that this is the case both when learning from static models [11, 12] as well as from dynamic visualizations of 2D or 3D models [13, 14]. In relation to CLT, Huk [13] found that while students with a high spatial ability benefited from interactive 3D models, those with low spatial ability did not. When asked, the high spatial ability students reported their (perceived) cognitive load to be low, whereas the opposite was true for low spatial ability students. This suggests that students with low spatial ability were cognitively overloaded by the presence of 3D interactive models, while high spatial ability students benefited from them as their cognitive load remained within working memory limits.

### *4.3. Phase Three (UG)*

The non-haptic group scored significantly higher than the control in one of the four question groups, while the haptic group did not score significantly higher on any occasion (despite scoring slightly higher than the non-haptic group on two occasions). This lends some support to the finding from the previous two phases. Student feedback again suggested a mixed response to the haptic interface with comments such as: *"(the haptic interface) slowed me down at first but it got better with time"* supporting the notion that for some at least there is an increased cognitive load imposed by the haptic device. As before there were also those who found the haptic interface to be beneficial: *"It made it far more enjoyable and engaging," "it improved my patience in dissection.* Some comments also suggested a feeling of missing out in not having access to the haptic interface: *"non-haptic was suitable for revision, but I feel haptic would be better for dissection practice."*

### *4.4. Phases Two and Three Combined (UG)*

When data from phases two and three was combined the non-haptic group scored significantly higher than the control in one of the four question groups, however the haptic group scored significantly higher than the control in all four of the question groups. It should be noted however that as sample size increases the statistical significance of an effect becomes greater. Both the control and haptic groups where combined (2014/15) to create larger samples of twenty seven and twenty respectively, compared with the non-haptic group (2015 only) of five. While this may go some way to explain the difference between these results and those of previous phases, it does highlight the limitations surrounding the use of small samples.

### *4.5. Conclusion*

The data gathered indicates that overall, those with access to the non-haptic version of the model performed equal or better than those with access to the haptic version. This is likely due to cognitive load being adversely affected by the addition of the haptic device. Students reported that the haptic device was not intuitive to use and took some time to get used to, if at all. No student used either version of the model for more than five hours, with over 40% using it for less than one hour. It is possible that with increased exposure to the haptic device students may find it easier and thus beneficial to use. The findings of this study however indicate that when used for a short period of time only (-5 hours) the haptic device may impede rather than enhance learning.

## References


[1] K. Leung et al, Anatomy instruction in medical schools: connecting the past and the future, *Advances in Health Sciences Education* **11** (2006), 209-215.

[2] F.R. Pryde; S.M. Black, Anatomy in Scotland: 20 years of change, *Scottish Medical Journal* **30** (2005), 96-98.

[3] L.J. Rizzolo; W.B. Stewart, Should we continue teaching anatomy by dissection when…? *The Anatomical Record* **289b** (2006), 215-218.

[4] J. McLachlan et al, Teaching anatomy without cadavers, *Medical Education* **38** (2004), 418-424.

[5] H. Hoffman; M. Murray, Anatomic VisualizeR: realizing the vision of a VR-based learning environment, *Studies in Health Technology and Informatics* **69 (**1999) 134-140

[6] B.M Ward et al, An evaluation of prototype VR medical training environment: applied surgical anatomy training for malignant breast disease, *Medicine Meets Virtual Reality 16* **132** (2008), 550-555

[7] S.N Biasutto et al, Teaching anatomy: cadavers vs. computers? *Annals of Anatomy* **188 (**2006), 187-190.

[8] A. Aziz et al, The human cadaver in the age of biomedical informatics, *The Anatomical Record* **269b** (2002), 20-32

[9] V.M Spitzer, The visible human dataset: the anatomical platform for human simulation, *The Anatomical Record* **253** (1998), 49-57.

[10] L. Verhoeven et al, Cognitive load in interactive knowledge construction, *Learning and Instruction* **19** (2009), 369-375.

[11] K. Rochford, Spatial learning disabilities and underachievement among university anatomy students, *Medical Education* **19** (1985), 13-26.

[12] R.S. Lufler et al, Effect of visual–spatial ability on medical students' performance in a gross anatomy course, *Anatomical Sciences Education* **5** (2012), 3-9.

[13] T. Huk, Who benefits from learning with 3D models? The case of spatial ability, *Journal of Computer Assisted Learning* **22** (2006), 392-404.

[14] S. Berney et al, How spatial abilities and dynamic visualizations interplay when learning functional anatomy with 3D anatomical models, *Anatomical Sciences Education* **8** (2015), 452-462.