The Design and Evaluation of a Haptic Veterinary Palpation Training Simulator

Andrew Crossan

Submitted for the degree of Doctor of Philosophy

University of Glasgow

Department of Computing Science

and Faculty of Veterinary Medicine

December 2003

Abstract

Virtual Reality (VR) simulators have the potential to offer a new training paradigm for improving the practical skills of novice doctors and veterinarians. They provide a training environment where novices can make mistakes and learn from these mistakes without consequences to the patient. However, if adoption of VR medical simulators is to become widespread it is essential that they are shown to provide benefit in areas that have previously been difficult or dangerous to train for. They must show benefit over existing methods, and more importantly show that they provide the expected training benefits to the user.

This thesis follows through the design and evaluation of a veterinary medical palpation simulator. It is structured around three main areas. Firstly, it examines the integration of a simulator into an existing veterinary medical course and demonstrates the benefits of using accepted computer aided learning design techniques. Secondly, it examines the problem of evaluation of a simulator. A thorough, structured evaluation is performed that examines different aspects of the training environment developed. Finally, methods of augmenting a simulation are discussed. Multimodal cues are introduced as a method of providing guidance and assessing the performance of users on the simulator. The case study discussed demonstrates the benefit of graphical playback cues for assessing performance.

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my supervisor Stephen Brewster, whose enthusiasm and insight made this thesis possible. Steve first introduced me to the area of haptics and has always provided encouragement, good advice, and good company since.

Thank you to Stuart Reid, and Dominic Mellor from Glasgow University Veterinary School whose guidance in all things veterinary was invaluable. Without them, I may never have appreciated the intricacies of large animal ovary palpation. Also, thanks must go to the many veterinarians and students from Glasgow University Veterinary School and Clyde Valley Veterinary Center. This list in particular includes but is not limited to Sandy Love, Monika Mihm, Kathryn Ellis, Karen MacEachen, Lindsay Robertson, Antony Clement, John Baillie, and John Maine. Last but not least, a special thank you to Sarah Baillie, who helped immeasurably throughout the project.

I would also like to acknowledge the Faraday Lab for help with the scanned ovary modelling, and in particular the technical wizardry of Colin McLaren and Ciaran Wills.

Thanks go to my officemates Ian Oakley, Marilyn McGee and Joy Goodman. The office discussions provided a heady mixture of entertainment and useful information that proved the inspiration for much of the work and helped me get through some of the longer days.

Thanks also go to my friends at Glasgow University and in particular the Multimodal Interaction Group for providing a wonderful and enjoyable atmosphere in which to work. Thank you to Peter Saffrey for the frequent discussions over coffee that ranged from the philosophical to the surreal, but always proved useful and above all entertaining.

Finally, thanks go to my family and in particular my parents for their encouragement and support.

Declaration of Originality

The material presented in this thesis is the result of my own research carried out at the Department of Computing Science at the University of Glasgow working under the supervision of Professor Stephen Brewster and Professor Stuart Reid of Glasgow University Veterinary School. All other referenced material has been given full acknowledgement in the text.

1 Introduction

1.1 Introduction & Background

There are inherent problems in providing training to novices in any safety critical task. In particular, teaching medical and veterinary medical clinicians to perform procedures poses challenges and possible risk to patients. Traditionally, these disciplines have relied on the apprenticeship model to train novice doctors and veterinarians. A novice will watch and provide assistance to an experienced medical practitioner performing a procedure several times before trying the procedure him or herself under the guidance of the expert. If the novice requires assistance or the expert judges that the novice might harm the patient in any way, then expert will step in to correct the actions of the novice or complete the procedure. The apprenticeship model has been in use to train doctors and veterinarians for over a hundred years [75]. Klein [59] describes the idea of this training system as follows.

> *"The basic, almost romantic notion is that the trainee learns the trade by doing, making minor errors along the way that are quickly recognised through the close observation of the teaching physician with 'instant' correction."*

There are obvious risks to the patient when the novice surgeon may not have the knowledge or the skills to perform the procedure. The training of new doctors and veterinarians is essential, however ethical considerations with regard to the welfare of the patient must be taken into consideration. The current system relies on the expert being a good teacher as well as skilled at the procedure. Also, apart from the danger to the patient posed by novice clinicians, there is no standardisation of method to perform the procedure. If the expert uses an outmoded or less successful technique to perform the procedure, then the novice will be taught that same technique.

Pilot training is often used as analogy to medical training. Novice aircraft pilots are similarly required to train in a complex and safety critical task. Pilots are required to react to unforeseen difficulties. Virtual Reality flight simulators for pilot training have grown in popularity since their initial introduction. As the technology has improved, and the fidelity of the simulations has increased, they have gradually become an essential part of pilot training for not only developing skills, but also maintaining skills [96]. The simulator allows the development of basic skills, but also allows pilots to train for emergency circumstances such as difficult weather conditions or mechanical failure. They can also be vital when training is prohibitively expensive such as training astronauts for missions in space.

Medical simulations are now widely viewed as a potential solution to many of the problems associated with medical training. Similarly to flight simulations, medical simulators offer the potential to provide training for a safety critical, multi-dimensional task. Simulation has been an accepted part of medical training for many years. Simulators allow novice clinicians to gain new skills by repeatedly practising in an environment where their inexperience and mistakes will not cause harm to a patient. Anatomy lab tutorials are a common method of providing training in anatomy and surgery skills, however scarce resources and large student numbers can restrict this form of training. Physical models, such as mannequins for teaching resuscitation are also available, but tend to be inflexible in their use. Students can only experience the cases that occur during their training, and therefore rare cases might have to be handled for the first time once qualified when the clinician will be expected to diagnose and treat the condition. Seasonal examinations - such as equine ovary examination - further restrict access to practical training. Simulators offer the possibility of allowing increased access by being available at any time during a course.

There are now several examples of advanced physical simulations that allow clinicians to monitor and practice skills on a wide variety of situations. Mannequins have been developed to allow doctors to practice ultrasound techniques on a model patient [1]. The University of Veterinary Medicine in Vienna have developed a fibreglass horse to simulate palpation of the intestine for colic.

One growing area of research is the use of Virtual Reality simulation as a teaching tool for medical education. With recent improvements and fall in the price of technology Virtual Reality systems are becoming more feasible as a training tool. Haptic devices are available that allow a user to interact with objects within a virtual environment, and 'feel' these interactions (haptics is a general term relating to the sense of touch [88]). Therefore, a haptic device is a device that a user interacts with using his or her sense of touch. Recent developments in haptic devices such as the PHANToM [72] from SensAble Technologies and the Impulse Engine [49] from Immersion have contributed by allowing natural interaction with the virtual environment and high resolution force feedback to be included into simulations, increasing the fidelity of the simulation.

One advantage of these virtual simulators is that a lot of flexibility can be built in. A large number of cases and conditions can be developed and run using the same equipment. This is particularly true when using a general purpose haptic device such as the PHANToM. Also, the anatomy and physiology can be integrated which is often difficult with plastic models. It can also be difficult to assess the performance of students on physical simulations.

Virtual Reality simulators have now been developed to simulate a wide variety of procedures and conditions. Minimally invasive surgery is the most prominent area where there are several examples of commercial and research systems [2, 14, 106, 113]. The rapid growth of minimally invasive surgery over the last ten years and the high dexterity required to operate the tools make this a promising area for simulation.

There are also several examples of surgery simulators for procedures that are not minimally invasive [5, 9, 42]. A palpation procedure involves the examination of a patient through direct contact. That is, using the clinician's sense of touch, and making the diagnosis through the feel of the area examined. Palpation simulators are also available for training in procedures that require the clinician to make a diagnosis through the haptic properties of tissue [16, 30]. However, these tend to be less commonly simulated possibly because in palpation procedures, the clinician is in direct contact with the patient.

There are now a growing number of companies now offering commercial Virtual Reality simulators for medical training. Immersion offers minimally invasive training systems, as well as Cathsim; a catheter insertion simulation. The Procedicus simulator from Mentice Medical Simulation [76] is designed to provide training for a variety of minimally invasive procedures. Novint Technologies [86] have developed a commercially available system for training in dentistry.

Until recently, the focus of this research has been mostly towards increasing the fidelity of the simulation. There has been a significant amount of research on making the simulation look and feel more like the real procedure. There are many examples of research examining the haptic properties of different types of tissue, and deformation of the virtual models when touched [9, 13, 63]. This is a non-trivial problem and the complex more accurate methods may require considerable computational power.

Only recently has there been an increase in research to show the usefulness of these systems in a teaching context. Showing that a training simulation provides benefit in the appropriate real world task is essential if a simulator is to gain widespread acceptance. Without strict validation, there is no way of knowing what benefit the users of the simulator are gaining, and if the actual skills trained match the skills supposedly trained. If a course relies on simulator training without strict validation, there is the possibly of under-trained doctors or veterinarians performing procedures that they do not have the necessary knowledge or skill to perform. This thesis is based on the integration and evaluation of a haptic medical simulator into a veterinary medical course.

1.2 Motivation

> *"In the course of development it must be realised that the expensive new technologies should bear the burden of proof of their effectiveness and reliability before they are put into training programs"* [8]

Simulators are often described as a risk free method of allowing novices to practice their skills. However, it can be argued that the risk to the patient still exists. It is merely shifted on to the design and validation of the simulator. A simulator does have the potential to improve safety during training in a safety critical task as demonstrated by the success and widespread adoption of flight simulators. It must be shown, however, that a simulator provides training in the stated tasks before it is integrated into a training course. If training benefits are not shown, the simulator may provide no benefit. This might have the effect of giving a novice surgeons or veterinarians a false sense of confidence that they can perform a procedure. In the worst case, a simulator might train users in such a way that it would be dangerous to perform the procedure similarly in real life. Validation of a simulator is therefore essential if the risk to the patient is to be reduced.

This however is not a trivial problem. There are obvious ethical concerns in performing a direct comparison in performance on the real task of novices trained using either traditional training or simulator training. A direct comparison would require that patients would be exposed to novice medical personnel whose only experience of the procedure is through a simulator without knowing if the simulator provides any benefits in the real world task. This would be exposing patients to unnecessary risks.

There have been some successes in this area. There are examples of simulators that have shown that repeated use can improve manual dexterity for and improve time taken and accuracy for targeting tasks using surgery and minimally invasive surgery tools [60]. This is an important result as it demonstrates that it is possible to gain skills using a simulator that will be useful in the medical domain. A study carried out by Gorman *et al.* [42] also indicates that this is the case for a venipuncture simulator. However, it does not necessarily follow that a user who performs well on a simulator will make a better clinician than one who does not. Gorman also notes that high manual dexterity demonstrated on a surgical simulator is not necessarily an indication of surgical skill. High manual dexterity is an important aspect of some procedures but there is a vast array of knowledge and technique that will also be required. Without being able to show that simulator training can provide similar benefits to traditional training, the advantages of simulators (such as flexibility and increased access) will mean little.

Many designers ignore the context of use of a simulator within a training course. The goal of the research is often to produce a high fidelity environment to simulate a given procedure. Many ignore the potential benefits that a virtual environment can bring to the training. VR training offers the possibility of providing an objective performance rating for any user. Currently, a doctor's ability to perform a procedure may be based on the number of times he or she has performed it. This may provide an inaccurate measure. By assuming properties for the virtual model, and monitoring the user's actions and forces used, it is possible to objectively rate his or her performance on metrics such as tissue damage. Supplying feedback on performance is important if users are to adjust their behaviour to improve performance.

1.3 Research Aims

In this section, the main research aims of this thesis will be discussed. This thesis will be examining the use of virtual environments, and in particular haptic environments, as a method for training veterinary students in palpation skills.

1.3.1 Thesis Statement

A haptic simulator can be used to provide a useful training environment for veterinary education. A simulator can provide features not possible using traditional methods to train and assess students such as multimodal feedback cues to augment the training.

1.3.2 Key Research Questions

To defend this thesis statement, the thesis will attempt to answer 3 key research questions:

1. Can a veterinary medical simulator be designed to integrate into a veterinary medical course to provide benefit in areas where traditional methods prove restrictive?
2. Can training through a medical simulator provide similar benefit to traditional training methods?
3. Can a medical simulator be used to allow a teacher to assess the performance of a student?

Question 1 asks how a simulator will be integrated into traditional courses. Current courses will all use various training aids such as books, lecture notes, and Web based material. What should the simulator simulate? Students may also learn through lecture and practical sessions. How will the inclusion of a simulator provide benefit to the students, and how will it affect other material in the course?

Question 2 asks about the benefits of simulator training. This is key to the inclusion of simulator training in the medical domain. In order for a simulator to become an accepted training tool, it must be shown to provide benefit in the area in which it has been designed to train. This is not a trivial task as ethical issues prevent direct comparisons.

Question 3 examines the benefits of using features available in a virtual simulation that would not be possible using traditional methods. Assessing performance is an important aspect of a training environment as this allows a student to discover his or her errors and attempt to correct them. This question examines whether features of a virtual simulation that are not present in traditional training can be used to provide assessment.

1.4 Thesis Structure

This chapter has provided an overview of the research in this thesis. It has provided a brief review of the background information for the problem area. Within this area, the key questions that will be addressed in the thesis have been identified. The following chapters will be structured around these questions, in order to attempt to answer them.

Chapter 2 is titled ‘Haptic Perception’. It contains an introduction to haptic perception and haptic devices. This chapter starts by defining important terms relating to haptic perception that are required for the thesis. The chapter contains information on the human haptic perceptual system, and how humans explore objects through their sense of touch. It is important for the thesis that the haptic technology is also introduced here. The technology imposes limitations on what can be simulated, and it is necessary to show that the devices that are available are sufficient for the task at hand. A brief description of some applications of haptic devices is also given.

Chapter 3 is titled 'Virtual Reality in Medicine'. This chapter contains the majority of the literature review for the thesis. The uses of Virtual Reality in medicine are briefly introduced before the topic of Virtual Reality simulation for medical training is described in more depth. The chapter describes current available research and commercial virtual simulator systems. The chapter covers systems that have proved successful in the field as well as examining limitations of the systems developed to date. It continues by describing some of the current research areas in medical simulation. The trade-off between complexity and fidelity is discussed. Validation of the simulator and providing performance feedback for the user are identified as important areas here. These two key areas are then explored in the later chapters in the thesis.

Chapter 4 is titled 'Design and Development of a Palpation Simulator'. It is included to investigate Question 1 of the key research questions from Section 1.3.2. This chapter initially introduces the ABC method for designing computer aided learning material. The following sections then describe how this method was used to design the Glasgow Horse Ovary Palpation Simulator (HOPS). It describes a series of interviews carried out with experts for requirements capture purposes, and discusses the integration of a simulator in the Glasgow Veterinary School course. The final section discusses the development of the Virtual Reality training system.

Chapter 5 is titled 'Validation of the Simulator'. This chapter is used to investigate Question 2 of the key research questions from Section 1.3.2. It describes three experiments carried out in an attempt to validate the simulator developed in Chapter 4. The first experiment involves the comparison of performance on the simulator over participants of different skill levels. This section describes an initial experiment that attempts to validate the HOPS simulator. Differences and similarities are discussed between the performance of experienced veterinarians and novice veterinary students on the simulator. Recorded position and force data from the experiment is also analysed in this section. The second experiment described examines the effect of multiple training sessions on the simulator on user performance. A group of novice participants were examined over four training sessions spaced one week apart, and a final training session one month after the fourth session. The first four sessions therefore look at the effect of regular training, where as the final session examines the retention of skills after a longer break. Changes in performance, time taken and workload data have been analysed and are discussed in this section. The third validation experiment builds on work in the previous experiment. The purpose of this is to compare Virtual Reality training with HOPS against traditional training methods. The performance on anatomy laboratory specimen ovaries of two groups of veterinary students is measured. The first group consists of the students from the previous multi-session training experiment. The second group consists of students who performed traditional ovary examination training. Workload results for both groups are also presented. Finally, conclusions are drawn from the experiments and their consequences for the validation of the HOPS simulator will be discussed here.

Chapter 6 is titled 'Augmenting Virtual Medical Training' and investigates Question 3 of the key research questions from Section 1.3.2. It initially contains a motivation for and a brief description of the concept of multimodal feedback cues. Several multimodal cues are described and methods are presented that could potentially be used to provide training and performance feedback to a user. Implementation details of these graphical, auditory and haptic cues are also presented here. This chapter describes an experiment to validate graphical cues as a tool for assessing performance on the simulator.

Chapter 7 is titled 'Conclusions'. This section contains conclusions reached from the work described in this thesis. References are made to the research questions stated in Section 1.3.2, and this section will discuss to what extent the questions can be answered by the work that has been carried out. Limitations of the work will be addressed at this point, and there is a discussion of issues to be resolved. This section will contain a list of issues with the work that were not resolved within the thesis, and make recommendations for future work to be carried out based on the conclusions found.

2 Haptic Perception

2.1 Introduction

This chapter will provide the definitions for terms used in this thesis as well as explaining the processes involved in haptic perception. It is important to know about the underlying processes involved, as without this it is difficult to understand the benefits and limitations of current haptic devices and the types of problems that they can be used to solve. It is also important to be aware of the techniques used in exploring a scene using our haptic sense. Current types of haptic devices and their strength and weaknesses will be introduced with some examples of the current technology. These will have an impact on the success of simulating different procedures.

2.2 Low Level Haptic Perception

The word haptic derives from the Greek word *haptesthai* meaning 'to touch' [99]. We perceive an object haptically through touching it. Touch is often referred to as a sense, but is more accurately a group of senses perceived through the skin. There are important differences between our sense of touch and our visual or auditory senses. To perceive an object through touch, contact is required with the object or some intermediary probe. Touch is also unique as a sense in that it is a fully duplex channel. As we contact an object, we apply a force on the object, but the object also applies a reaction force on us. In this situation, a feedback loop is formed such that we can affect the position or state of an object through contact, and adapt our actions to affect the object differently using information gathered through our sense of touch.

Important distinctions exist between different forms of touch. This section will provide an overview of the definitions used in this thesis for these different forms of touch. Although the terms referring to haptic interaction are commonly used in the haptic literature, problems arise as they are often used with slightly different meanings [15, 70, 112]. The definitions of the terms in this thesis will be based on those stated by Oakley *et al.* [88] shown in Table 1. These definitions have been synthesised from both the psychology and computer science communities, and provide clearly bounded terms based on the properties of the human haptic system.

Term	Definition
Haptic	Relating to the sense of touch.
Proprioceptive	Relating to sensory information about the state of the body (including cutaneous, kinesthetic, and vestibular sensations).
Vestibular	Pertaining to the perception of the head position, acceleration, and deceleration
Kinesthetic	Meaning the feeling of motion. Relating to sensations originating in muscles, tendons, and joints
Cutaneous	Pertaining to the skin itself or the skin as a sense organ. Includes the sensation of pressure, temperature, and pain
Tactile	Pertaining to the cutaneous sense, but more specifically the sensation of pressure rather than temperature or pain
Force Feedback	Relating to the mechanical production of information sensed by the human kinesthetic system

Table 1. Definitions relating to touch stated by Oakley *et al.* [88]

The human haptic system *is "the entire sensory, motor and cognitive components of the body brain system"* [88]. Proprioceptive is defined as *"Relating to sensory information about the state of the body (including cutaneous, kinesthetic, and vestibular sensations)"*. The term haptic is therefore closely linked to, but contained within, the definition of proprioception. The most important distinction for the purposes of this thesis however, is between the terms cutaneous and kinesthetic. It is also important to realise that the distinction has not been arbitrarily chosen. The human body senses each of these through different mechanisms. This distinction between the cutaneous and kinesthetic systems becomes important when describing types of haptic devices and their uses and limitations.

Cutaneous perception occurs through stimulation of receptors that exist in the outer layers of the skin. Pain, temperature and tactile perception make up our cutaneous sense, but are sensed using different receptors in the skin. The terms tactile and cutaneous are often used interchangeably [15, 112], although this is somewhat incompatible with the above definitions. Tactile information is gathered from different types of receptors that exist in the outer layers of the skin. These different types of receptor respond well to different frequencies of vibration. Low frequencies would correspond to pressure on, or stretching of, the skin, where as high frequencies would correspond to a vibrating or buzzing sensation on the skin. The human cutaneous system can perceive vibrations in the range of 0.4Hz to over 500Hz [41]. The upper limit of 500Hz is an important consideration when considering simulation of a solid object rather than simulation of a vibration [78]. This is discussed further in Section 2.4.

Some areas of the skin are more sensitive to contact than others, but a person can normally perceive an object haptically through contact with any point on his or her skin. In fact, the resolution of tactile perception varies widely throughout the human body, and increases as the density of receptors at the stimulated area of skin increases. The fingertips are one of the highest resolution areas, which is to be expected since this is the common area humans use to explore objects haptically. A common method of measuring this resolution is using the 'Two Point Threshold' method which is *"the smallest separation between two points on the skin that is perceived as two points"*, and can be as low as 2 or 3 mm at the fingertips [41]. Johnson *et al.* however [53] suggest that this method - although still in common use - is flawed. They present evidence from several studies that suggest that measured two point thresholds can vary quite dramatically even when measuring the same subject on a day to day basis. They suggest a method of measuring spatial discrimination thresholds using square wave gratings that has been shown to provide consistent results. Two sets of orthogonal gratings of varying frequency are used in the study. The skin area is pressed against a grating and the participant must judge its orientation. The threshold value is given as the gap width at which performance was achieved that was half way between perfect recognition and chance (75% for two orientations). The results presented suggest a spatial acuity of 0.94mm at the finger tips, and a slow as 0.51mm for the tongue.

Temperature and pain information, which also make up the cutaneous sense, are sensed separately from pressure through thermoreceptors and nociceptors respectively [15]. Thermoreceptors respond to a change in skin temperature. Again, the sensitivity of the area of skin to changes in temperature depends on the density of receptors. Nociceptors will respond to tissue damaging stimuli or over stimulation of any other receptors. This may be the result of harmful chemicals being placed on the skin surface, or highly intense mechanical or thermal stimuli. Although the pain stimuli may be applied to the skin for a short period of time, the pain response may continue for some period of time.

Kinesthetic perception is closely linked to proprioception, which allows us to sense the current position of our limbs. Lindermen and Templeman [70] use these terms interchangeably. The definition provided by Oakley *et al.* [88] groups cutaneous, kinesthetic and vestibular senses under the term proprioception, and this is the definition used throughout this thesis. Both are sensed through stimulation of receptors in the muscles and joints that are sensitive to changes in length of the relevant muscles. We perceive an object kinesthetically through movement. When a person explores an object using his or her kinesthetic sense, it is most often through limb movement. A person will move a limb to contact an object, and will perceive a reaction force from the object that restricts the movement of the limb. We perceive an object kinesthetically through this reaction force. Most often when exploring an object, a person will receive information through his or her kinesthetic and cutaneous channels together, and combine this information to form a coherent model of the object.

The kinesthetic system would only be in use when an object is being actively explored. It is therefore important to make the distinction between different modes of touch. Active touch refers to a person touching a static object [41]. A person is actively exploring the object him or herself through limb movement, and is providing the movement and force that stimulate the receptors. Passive touch is the other extreme, and can be described as a sensation on the skin [41]. A person would remain stationary as the object was moved over his or her skin. In this case, it is the movement of the object that provides the stimuli. As kinesthetic perception requires movement of part of the body, it can therefore be seen that the kinesthetic system is not used in passive touch. The modes of touch provide different sensations to a person, even when the same object is explored by both methods. Although relative motion of the object exists in both cases, a person actively exploring an object will perceive it as stationary. A passively explored object is however perceived as moving by the stationary observer [41]. The modes of touch described above are the extreme situations. Situations also exist where a person will actively explore a moving object. In this situation, the stimulation is due to both a person's movements, and the movements of the contacted object.

A situation might also occur where a person is guided through the exploration of an object. This would mean that a person who moves along a path is guided or restrained by an external force when exploring a static object. This is very similar to active touch in that the subject is moving while the object can remain stationary, so the kinesthetic sense is still used in the exploration. The subject does not control his or her motion, and can therefore be seen to be passively guided. This is similar to Smith's [108] definition of passive touch, although it is clearly conflicts with the generally accepted definition of passive touch [41].

Carello and Turvey [19] also introduce the concept of dynamic touch through cues from a wielded object. A wielded object will deform tissue in the hand, arm and body. These deformations will provide information to the wielder about it's weight and distribution of weight. Moving the object will lead to changes in these deformations due to the object's inertia that will allow the wielder to perceive information about the object. Carello and Turvey describe how dynamic touch can be used to judge the length of a cane and how a subject's perception of the length of the cane could be influence by placing weights closer to or further away from the position held and therefore changing the distribution of weight of the object.

2.3 Haptic Exploration in the Real World

Haptic exploration is the term used to describe exploration of an object through touch. When a person is exploring an object, it is most often done to gain information to identify the object or some feature of the object.

2.3.1 Haptic Object Identification

In many situations when examining an object, we use more than one point of contact. In some instances, a person will use both hands, and several fingers at the same time to build up a mental representation of an object. The human haptic system can take the separate information gathered from different parts of the body, and combine the haptic information to build up a coherent model of the object. This is the natural way in which humans explore objects haptically. Studies have shown high levels of accuracy for object recognition using the haptic system.

A study conducted by Klatzky *et al.* [58] examined haptic object exploration with no constraints. Twenty participants explored one hundred common objects using touch alone. Participants were blindfolded and allowed to explore each object with both hands. They could also pick up the object if required. A timer was started when the participant first contacted the object. The timer was stopped when a microphone detected noise from the participant verbally naming the object. A visual identification task was initially carried out to ensure that the objects chosen were consistently nameable. Five separate participants identified the objects when shown them. The 100 objects chosen for the study were identified correctly by all five participants. Of the 2000 objects haptically explored, 83 were incorrectly identified. Of the incorrect responses, 22 were classed as superordinate errors, where the answer returned was too general. For example, the participant responding 'vegetable' instead of 'pumpkin'. A further 14 were superordinate errors that were then corrected by the participant. 29 incorrect responses were categorically related. That is, items were of the same type, such as 'sock' instead of 'T-shirt' which are both items are clothing. 14 errors were responses unrelated to the object, and 4 errors were omissions. When applying a strict naming convention, 94% of responses were correct. With a less strict convention allowing related responses, 99% accuracy was achieved. 68% of responses occurred within 3 seconds, and only 6% took longer than 6 seconds. These results show that the human haptic system is extremely successful at quickly identifying familiar objects.

2.3.2 The Haptic Glance

The visual glance gives a person an immediate overview of an object. Klatzky and Lederman [57] introduce the idea of a haptic glance. During a haptic glance, severe spatial and temporal constraints are placed on exploration. Klatzky and Lederman constrained participants to either 200ms or 3000ms during the haptic glance. Participants were not allowed to move their fingers over the surface of an object. Klatzky and Lederman demonstrated that a recognition rate of above 20% was possible using these conditions. In this study, participants used finger position and information passed through the tips of their fingers to identify the objects using structural and material properties such as global shape and roughness.

A control group from this haptic glance experiment managed an identification rate of 93% when no spatial or temporal constraints were placed on their exploration. This is consistent with the results presented by Klatzky *et al.* [58] in 2.3.1.

2.3.3 Modes of Touch

Active touch is traditionally thought to provide a better level of recognition for haptic exploration. Gibson demonstrates this experimentally [40]. Cookie cutters of different shapes with a mean diameter of approximately 2.5cm were presented to participants for them to explore using touch alone. A curtain was placed between the participants and the cookie cutters to ensure that the objects could not be identified visually. Participants who were allowed to actively explore the cookie cutters correctly identified the shape 95% of the time. The passive case consisted of cookie cutters being pressed into a participant's hand with the participant correctly identifying the shape 48% of the time. However, an extra passive case showed that passive touch can be just as effective as active touch in haptic identification. In this case, participants felt the shapes as the experimenter moved cookie cutters over their fingers. The participants did not perform significantly worse than in the active touch case. Gibson concludes that the important factor in recognition is the lateral relative motion of the object across the skin.

2.3.4 Exploratory Procedures

Lederman and Klatzky experimentally identified a number of exploratory procedures (EPs) that a participant uses when discerning different properties of an object or surface [66]. They describe an EP as a *"stereotyped movement pattern having certain characteristics that are invariant and others that are highly typical"*. Participants were asked to explore a series of objects through touch alone, and asked to identify specific properties of the object such as surface texture, temperature or weight of the object. The movements of these participants were recorded and analysed *post hoc*, and it was noted that highly typical gestures were used to identify the different properties of the objects. Figure 1 shows the set of EPs that Lederman and Klatzky identified. Each of these procedures can be associated with determining specific properties of the object.

- 'Lateral Motion' is used to determine texture, and involves the sideways movement between the object and the skin, typically on an interior surface rather than an edge.
- 'Pressure' is used to determine hardness. The object is often fixed either by the subject's hand or another external force, while he or she applies normal forces to the object's surface.
- 'Static Contact' occurs when the object is supported (by an external force or the subject's other hand) while one hand passively rests on the objects surface. This is used to determine temperature.

- 'Unsupported Holding' can be used to determine an object's weight. This involves supporting the object away from any external surface without moulding the hand to the object, and could also involve hefting of the arm or wrist.
- 'Enclosure' is typically used to determine the global shape, or volume of an object. The participant tries to envelop as much of the object as possible in an effort to mould his or her hand to it's shape. The subject may shift the object within his or her hands.
- 'Contour Following' can be used to trace the exact shape of an object or determine the volume. The subject's hand maintains contact with the contour of the object typically using smooth non-repetitive motions.
- The 'Function Test' EP involves executing movements that perform certain functions of the object, for example pinching the end of a set of pincers. This is used to determine a specific function of the object.
- The 'Part Motion Test' is used to identify moveable parts of an object. The participant will apply a force to the specific part, while applying a counter force to stabilize the rest of the object.

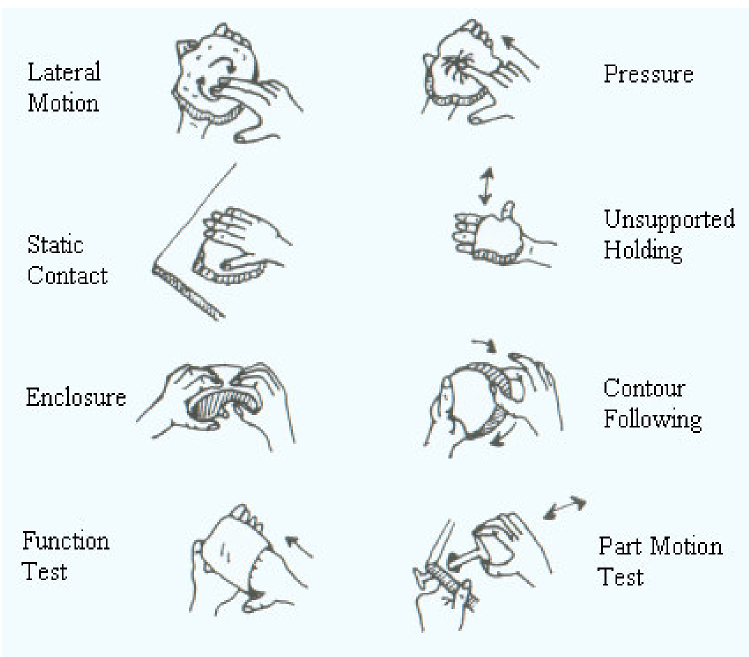


Figure 1. Typical exploratory procedures described by Lederman and Klatzky for determining object properties through touch.

In Table 2, Lederman and Klatzky [67] indicate the relative precision of some of the EPs for extracting information about the different properties of an object. This is indicative of the fact that more than one EP can be used to identify a property, however, some EPs are better at identifying a property than others. A higher number indicates a higher precision that can be achieved in identifying that property of the object using an EP. In the event of similar accuracy, a higher number is given to the EP that can be used to identify the feature more quickly.

	Texture	Hardness	Temperature	Weight	Volume	Global Shape	Exact Shape
Lateral Motion	2	1	1	0	0	0	0
Pressure	1	2	1	0	0	0	0
Static contact	1	0	2	0	1	1	0
Unsupported Holding	0	1	1	2	1	1	0
Enclosure	1	1	1	1	2	2	0
Contour Following	1	1	1	1	1	1	3

Table 2. EP to property weightings. A higher value indicates a higher relevance of the EP to haptically identifying the selected property of an object.

These exploratory procedures can be used to determine the limitations of haptic devices (see Section 2.4) as each of the current devices only support a subset of these EPs. They can also prove useful in the medical domain when analysing a palpation procedure to determine the essential motions. The previous research described in this section shows that the human haptic system can identify objects quickly and accurately in the real world. Section 2.5.1 describes haptic object identification in the virtual world and compares this to real world performance.

2.4 Computer Haptics

Cugini *et al.* [23] describes a haptic device as:

> *"any output device whose interface with the user occurs through its motor and/or tactile system"*

A haptic device is often combined with position sensors to provide a mechanism for input, and hence allow a fully duplex input and output device. They are often used as a means of interacting with a virtual environment, and can be used to indicate a contact or constrain the user's movements to simulate interaction with objects or effects within the scene. The benefit of this can be increased realism in the scene by incorporating the haptic sense into the simulation. Without the benefit of a haptic device, a user may find it difficult to determine when he or she is touching a virtual object. Haptic feedback can indicate to the user when an object is touched, grasped, or is slipping for example.

There are several different methods of classifying haptic devices that will be discussed in this section. The most obvious classification is in the type of haptic feedback being provided by the device.

Current devices can generally be split into either devices that provide tactile feedback (through the cutaneous sensory channel) or devices that provide force feedback (through the kinesthetic sensory channel). There are however some examples of devices where a combination of different forms of feedback are presented to the user [90, 117].

2.4.1 Tactile Devices

Tactile devices provide tactile stimuli to users, to present them with information. They typically consist of one or more effectors that transmit tactile stimuli to the skin. Shimoga [107] presents four methods that have been used to provide these stimuli; Pneumatic, Vibrotactile, Electrotactile, and Functional Neuromuscular stimuli. A further method of providing tactile stimuli by raising and lowering blunt pins will also be considered.

2.4.1.1 Pneumatic Stimuli

Shimoga describes three different methods of providing pneumatic stimuli to the user. Air jets fired onto the skin can be used to indicate a contact. An array of jets fired onto and area of skin can be used to provide a higher resolution and present the user with patterns. However constant exposure to air jets can lead to temporary numbness, and loss of tactile abilities. Minute air pockets can be placed against the skin. A contact would be indicated to the user by inflating the air pocket such that the inflated pocket presses against the skin. Alternatively, air rings can be placed around the user's fingers and inflated as the user contacts an object. User fatigue becomes an important issue when inflating pockets or rings against the skin for long periods of time. Device response times must also be taken into consideration as the tactile experience will be affected if the air pockets or rings take too long to inflate or deflate.

The Teletact II device [111] (shown in Figure 2) developed by the Advanced Robotics Research Device Centre is an example of a device that supplies pneumatic stimuli. The user grasps the device, and tactile feedback is presented to the user using thirty small air pockets.

Figure 2. The Teletact II glove.

2.4.1.2 Vibrotactile Stimuli

Vibrotactile stimuli are provided by one or more vibrating effectors on the skin. These are commonly seen as an array of blunt pins, similar to a Braille device. Vibrating the effectors at a specific frequency can be used to indicate a contact. A user would feel a tickling sensation on his or her skin that can be interpreted as a contact with a virtual object.

Figure 3. The Cybertouch tactile glove devloped by Immersion.

The Cybertouch device [47] (shown in Figure 3) available from Immersion is one example of a vibrotactile device. This device is often combined with a position sensing device, that allows a user to manipulate an object in three dimensional space. Tactile feedback is presented to the user through 6 vibrotactile stimulators; one on each finger and one on the palm. Once a user contacts an object, there is no method for restricting his or her movements so that the object is not penetrated. The tactile stimulators can however indicate that part of the the user's hand is in contact with an object.

2.4.1.3 Electrotactile stimuli

Electrotactile stimuli occur by introducing electric pulses to the skin of the user to indicate contact with an object. Careful choices need to be made about the electric pulse width and frequency so as not to cause the user pain. Different areas of the skin respond better to different frequencies of stimulation, so placement of the effectors is important and may depend on the frequency of stimulation.

Kajimoto *et al.* [54] state that one advantage of the use of electrotactile displays over vibrotactile arrays is that they avoid mechanical difficulties. However, they note the difficulties in confining the sensation to a small, focussed area. They also note the problem that electrical stimulation can feel invasive, and provoke a sensation of fear in users. Although users will not be harmed by a small electrical current, they can often be shocked by a sudden sensation not due to increased pressure on the fingertip, but just from touching the source. Kajimoto *et al.* [54] also present an electrotactile mouse shown in Figure 4. Stimulation is provided to the user through a 4 x 4 array of electrotactile stimulators mounted on the middle mouse button. By adjusting the pulse amplitude and width of the current, they can set different feels for an object. However, they note that presenting electrotactile stimulation that feels more natural is still an open research topic.

Figure 4. An electrotactile mouse developed by Kajimoto et al. [54].

2.4.1.4 Functional Neuromuscular Stimuli

Functional neuromuscular effectors are different from those described above as they use electric pulses to directly stimulate the somatosensory cortex in the brain. This raises ethical issues as functional neuromuscular stimulation is an invasive technique, with the possibility of damaging the user. As such no such tactile devices exist, as yet.

2.4.1.5 Raised Blunt Pin

One common method of providing tactile feedback to users not addressed by Shimoga is the use of an array blunt pins that can be raised and lowered. This method uses pins similarly to the vibrotactile feedback method. However, information is passed to the user through the pattern formed by the pins set at different levels. A user often uses his or her fingertips to detect patterns formed by the array, and these patterns are used to convey information to the user. The Powerbraille display [37] (shown in Figure 5) distributed by Freedom Scientific is a typical example of a device that conveys tactile feedback to a user through such a method. The two different models consist of 40 or 80 tactile arrays of blunt pins with 8 pins available for each cell. Each cell can be used to display one Braille character, and a visually impaired user can move his or her fingers tips over the cells to read the different characters.

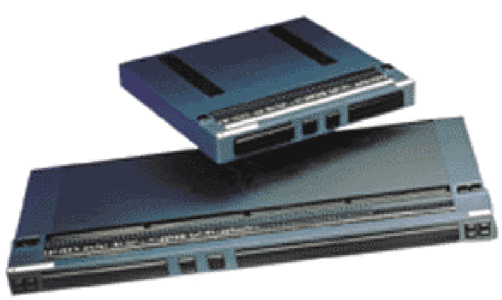


Figure 5. The Powerbraille braille device.

The VTPlayer mouse developed by VirTouch [116] (shown in Figure 6) is a further example of a raised and lowered pin system. It contains two arrays of 16 pins each. A user actively moves the mouse over objects in specially designed applications and the raised pins form patterns that may represent characters, textures, or the outline of an object. In this system, it is up to the user to actively explore the environment. This gives the device more flexibility than standard systems for displaying non-textual objects such as shapes or textures. This could however lead to problems if the device is used to display braille characters. If the mouse cursor is not directly over the centre of the character to be displayed, the Braille character displayed by the device will be displaced on the cell array and may not be fully displayed. Using systems such as the VTPlayer mouse, the user is restricted to two dimensional object exploration only due to the two degree of movement of the mouse.

Figure 6. The VTPlayer mouse developed by VirTouch. The right hand image shows the two arrays of blunt pins.

2.4.1.6 Applications and Limitations

Often, tactile devices are used as indicators to convey extra information rather than to haptically identify objects. Enriquez *et al.* [35] present a system that uses air pockets on a steering wheel to provide drivers with alerts. They were able to show a significant improvement in driver response times when their tactile steering wheel was used. Van Veen and van Erp [115] present a study to investigate user response times to tactile stimuli under high G-force. This is with regard to presenting fighter pilots with directional information through vibrotactile stimulators attached to the left or right side of the torso. This study showed that participants responded with a high degree of accuracy - between 85% and 100% - and response times were stable at around 500ms. They conclude that the tactile channel could be useful for displaying information to pilots.

One failing of tactile devices is that most are built for sensory substitution. Although they can provide a stimulus to indicate a contact, the stimulus might not be representative of the object being touched and the user will be required to interpret it. Also, the user's movements are not constrained, since there is no mechanism for applying large scale forces to the user's limbs through a purely tactile device. For this reason, virtual objects cannot have weight or rigidity. Unsupported holding, and pressure EPs require force to be applied to the user so cannot be supported by tactile devices. For these reasons, force feedback devices have largely replaced tactile devices for commercial and research projects.

2.4.2 Force Feedback Devices

Force feedback devices provide a mechanism to apply forces to a user. These forces can be structured to provide resistance and constrain a user's movements, which can present the user with the illusion of touching a physical object. However, a force feedback device may also have the functionality to produce forces to drive the user's motions. A user will interact with the device through his or her kinesthetic sense. Most often, a force feedback device will sense the position of the user in some environment, and set the force that the user feels based on where the interaction point is in the environment. Because of the high bandwidth of the kinesthetic channel, the update rate required to present a convincing illusion of touching a solid object needs to be relatively high. Minsky *et al.* [78] noted that the human haptic channel can sense vibrations up to approximately 500Hz. It is therefore necessary to exceed this update value when providing force feedback. Stability issues also affect the haptic update rate. A device may become unstable when a large change in force occurs in relatively few force updates. This may become an issue simulating very rigid objects. This also means that when communication bandwidth is limited (for example when communicating over the Internet), and a user is interacting with soft objects only, the update rate can be lower. Cugini *et al.* [23] present several mechanisms for classifying force feedback devices.

2.4.2.1 Passive and Active Devices

Passive devices supply force to the user through the use of brakes. Contact with a virtual object or effect, such as friction, can be simulated by presenting the user with arbitrary levels of resistance to movement. Passive devices are limited in comparison to active devices. Effects such as the release of a compressed spring cannot be modelled, as a passive device cannot be used to drive a user's motions.

Active devices use motors to provide forces to the user. Forces produced by these motors can be used to provide resistance to movement as in a passive device, but can also be used to actively drive the user along a path. This provides a greater range of situations that can be simulated. Although active devices can provide more functionality, they also tend to be more expensive to build. They generally require higher update rates, which may be important in situations where computational power is limited.

2.4.2.2 Number of Degrees of Freedom

The number of input degrees of freedom (DOF) is equal to the number of dimensions of movement that the device can allow, and can sense the position or angle of. The number of output DOF is equal to the number of dimensions of movements that have actuators associated with them, and therefore can provide force feedback to. The number of input and output DOF provided by a device is not necessarily the same. For example, the standard PHANToM haptic interface from SensAble

Technologies [72] allows six input DOF (translational and rotational), but only three output DOF (translational).

2.4.2.3 Points of Contact

This represents the number of points of interaction that the user has with the virtual world. A device with five points of contact will allow a user to simultaneously contact a virtual world in five places. Most common force feedback devices are restricted to one point of contact. This is due to the complexities and expense involved in adding more. A device such the PHANToM [72] provides one point of contact but provides very high resolution force feedback. More than one device can be used in a virtual environment however, allowing for more than one point of contact. Current force feedback devices do not present the user with tactile feedback distributed over the skin. Therefore, when using a force feedback device, a user can often be thought of as interacting through an infinitely small point of contact. This means that when exploring a virtual world through a force feedback device, the user will not be able to rely on cues distributed through the skin to provide information. The result of this is that it is difficult to perceive localised features or textural information from a virtual object without moving the contact point over the surface of the object. The results from Klatzky and Lederman's 'haptic glance' experiment discussed in Section 2.3 suggest that during real world exploration, the human haptic system could identify familiar objects around 20% of the time through these cues alone with severe time constraints in place.

A one point of contact device cannot not support the EP enclosure, where as many multipoint devices support a limited form of enclosure. For example, glove based force feedback devices often present one DOF to each of the fingers that allows a user to enclose a virtual object. However, current force feedback devices still present a user with distinct points of contact on the skin and not the continuous skin contact that would be felt during enclosure in the real world. When a user is restricted to one point of contact, exploration of an object becomes a temporal task as well as spatial. The user must move over the objects surface to determine shape and texture properties of the object.

2.4.2.4 Application Domain

A general purpose haptic device should be suitable for any haptic application, although this is not the case in practice. Kinesthetic receptors are spread throughout the skin and are not localised to one area so a general purpose haptic device allowing interaction with different body parts is difficult to envisage. The PHANToM [72] is an example of a device that attempts to be general purpose. It allows point interaction with a three dimensional environment, while providing the option of attaching customisable gimbals for different forms of interaction. Application specific devices are designed for one particular purpose or task. An example would be Immersion's Laparoscopic Impulse Engine [48] designed to simulate minimally invasive laparoscopic surgical procedures.

2.4.2.5 Grounded or Ungrounded Haptic Devices

On top of Cugini's categorisations above, another important factor is whether the device is grounded or ungrounded. When a haptic device applies a force to a user, this force must be balanced by an equal but opposite reaction force. Richard and Cutkosky [95] describe the differences between grounded and ungrounded devices. For grounded devices, this force is supplied by a large object such as a desk or wall. These devices can be described as being 'fixed' to the ground. Ungrounded devices can also be described as 'user grounded' devices. It is the user that supplies the force to counteract the reaction force from the device.

Grounded devices have the ability to provide force that will restrict the limb movements of a user. They have been shown to be successful in presenting a user with the haptic properties of virtual objects such as shape, size, stiffness and texture [51, 74, 118]. The main advantage of ungrounded devices is their portability. They tend therefore to have a larger workspace than grounded devices. However, there has been little research into the success of displaying a virtual object using an ungrounded device. There can be a certain degree of ambiguity involved with virtual objects displayed through an ungrounded device. For example, with a force feedback glove such as the Rutgers Hand Master (shown in Figure 7) a button push can be simulated by applying a force to one finger of the user. However, there is nothing to restrict the user from pushing his or her hand against the virtual button. In the real world, this would also have the effect of pushing the button but there is no way to recreate this reaction force using this device.

2.4.3 Criteria for an Effective Force Feedback Interface

Massie and Salisbury [72] identified three necessary criteria for an effective force feedback interface.

- *"Free space must feel free"*
- *"Solid objects must feel stiff"*
- "*Virtual constraints must not be easily saturated"*

When a user is navigating through the environment without touching an object, there should be no noticeable force to restrict his or her movements. There will always be some friction due to moving part in the device, but this should be minimised. There should be no weight imbalance such that the user needs to support the end effector of the device. The device should be able to provide a convincing simulation of a stiff object, although no force feedback device will be able to simulate a perfectly stiff object. Also, a user must not be able to overpower the device too easily. This will depend on the type of interaction taking place. The device must provide an appropriately high maximum force for its purpose. A user interacting with a device through his or her finger may rarely apply more than 10 Newtons of force. However, a device specifically designed to simulate a surgical procedure such as cutting bone, may be required to exert higher forces.

2.5 Haptic Exploration Through Force Feedback Devices

2.5.1 Virtual Reality Haptic Object Perception

Section 2.3 examined object identification through touch in the real world. This section will examine if haptic object identification is possible in the virtual world. There are differences to haptic exploration in the real world and in the virtual world. One major difference is in the type of feedback being supplied to the user. The human haptic sense is a bundle of senses that are integrated during exploration to give an impression of the haptic properties of an object. Most haptic devices do not support more than one type of haptic feedback. For example, it is rare for a force feedback device to provide cutaneous feedback to the user although there are examples. Wall presents a device that is attachable to a force feedback device that allows tactile information to be displayed to the user [117]. Ottensmeyer and Salisbury have developed a device that is attachable to a force feedback device, and can present the user with temperature information [90].

The main disadvantage that virtual haptic world exploration has over real world haptic exploration is in the restricted number of contact points. The skin contains receptors throughout the body that sense contact and provide feedback to the brain. Even the most complex force feedback devices are limited to few points of contact. Jansson [52] introduces the concept of perceptual filling-in for haptic exploration. He states that in the visual sense, two moving dots are often perceived as the ends of a rod. The brain fills in the gaps between the dots. There is currently no research that explicitly states that the human haptic system can perform this also. There have however been a number of studies involving participants who could successfully identify virtual objects through a haptic device. Jansson presents a study in which two groups of participants were asked to explore a series of simplified virtual faces through touch alone. The two participant groups were sighted participants and visually impaired participants. All participants used a PHANToM [72] - a one point of contact force feedback device - to try to identify features of the faces. The task was to state which of the facial features presented deviated from normal size and whether it was too large or too small. The features used were eyebrows, eyes, ears, mouth nose, and chin. The level of complexity of the face was controlled by controlling the number of features present on the face. In the simplest case (one feature), participants were able to determine the correct case approximately 80% of the time. This level of accuracy was achieved in a mean time of approximately 60 seconds for sighted participants, and 45 seconds for visually impaired participants. In the most complex case when all six features were present, the mean accuracy was reduced to under 40% and mean time taken for the task had risen to approximately 80 seconds for the sighted group and 60 seconds for the visually impaired group.

When compared to Klatzky and Lederman's results discussed in Section 2.3, it is clear that virtual exploration using a one point of contact force feedback device is far less successful than real world haptic exploration. Haptic exploration of the virtual objects took far longer than the real world exploration presented by Klatzky and Lederman. What is important to note though is that identification of virtual objects is still possible, even when users are restricted to one point of contact and kinesthetic information only. Jansson's results show that participants could achieve an accuracy of approximately 80% in the simplest condition. While this is less than would be expected in real world exploration, it shows that it is possible to achieve a high degree of accuracy when identifying virtual objects.

2.5.2 Haptic Devices

When choosing a haptic device for use with a specific application, it is necessary to take into account the types of exploration required for the task. Lederman and Klatzky's Exploratory Procedures have been discussed in Section 2.3. Current haptic devices only support a subset of these EPs, so the haptic device chosen must support as many EPs suitable for the task as possible. For example, grounded haptic devices will support a different subset of EPs than ungrounded devices. The following section reviews some of the currently available devices, and looks at their strengths and weaknesses for certain tasks.

2.5.2.1 The Rutgers Hand Master II

The Rutgers Hand Master II, shown in Figure 7, [11] is an ungrounded glove based force feedback device that allows one degree of force feedback each on three fingers and the thumb of one hand to restrict closing the fingers. This device will therefore support the EP pressure. Also due to the multiple contact points attached to different fingers, a limited form of enclosure is possible. This is only limited enclosure as feedback is received through three effectors only, and not through continuous contact with the skin. As only one degree of freedom exists for each point of contact, lateral motions and contour following will not be supported by the Rutgers Hand Master II. No force feedback is available on the arm when using this device, only on the fingers. It is therefore possible to move through objects in a virtual scene, and unsupported holding is not supported by the device.

Figure 7. The Rutgers Hand Master II force feedback device.

2.5.2.2 The PHANToM

The PHANToM from SensAble Technologies [72], shown in Figure 8, is the most widely used force feedback device for research applications. It is a grounded force feedback device that allows a user to move freely in 6 degrees of freedom (x, y, z, roll, pitch and yaw). The device can also provide 3 degrees of high-resolution active force feedback to resist or assist motion in the x, y and z dimensions. A user interacts with the PHANToM using a gimbal attached to the end of a mechanical arm.

Figure 8. The PHANToM force feedback device from SensAble Technologies.

The standard gimbals for the device are a thimble (as shown in Figure 8) and pen, although different end effectors can be built and attached to the end of the PHANToM arm if required. Position sensors are present on the pen attachment to allow position sensing for roll, pitch and yaw. The device provides a maximum of approximately 10 Newtons of force for short periods, and approximately 1.4 Newtons of continuous force. This is sufficient for most interactions through the fingertip. The PHANToM allows interaction with a virtual environment through a single infinitely small point of contact. This limitation is due to the fact that a cursor with a fixed size may be required to prevent a user from rotating through a virtual object. However, the standard PHANToM models do not provide force feedback in the rotational dimensions. The PHANToM 6DOF [104] does provide rotational force feedback but is currently not in widespread use. There is the option to use two or more PHANToMs in the same environment, and therefore allow more than one point of contact interactions. This would allow a user to grasp objects in a virtual scene.

Wall and Harwin [119] provide an overview of the EPs supported by use of a single PHANToM in an environment. They conclude that lateral motion, static contact, contour following, pressure and unsupported holding are provided by the device. Static contact is only supported in a limited form as temperature information cannot be conveyed through the PHANToM. Contour following is limited with the PHANToM as tactile cues are not provided to gain information about local features such as contours. Enclosure is however not supported, as the PHANToM is a one point-of-contact device.

By using two or more PHANToM devices in the same environment, it could be argued that a limited form of enclosure is supported. This would be enclosing an object using discreet points on the surface, which will feel very different to the continuous skin contact felt during enclosure in the real world.

By using the standard GHOST toolkit, geometric and VRML objects can be incorporated into a PHANToM environment to provide the touchable model. For each of these models, stiffness, friction and damping properties can be set to provide a variety of different feels.

2.5.2.3 The HapticMASTER

The HapticMASTER [114] developed by FCS Robotics is shown in Figure 9. It uses a similar mechanical arm structure to the PHANToM to provide force feedback. One of the main advantages of the device over the PHANToM is that the maximum forces that the device can provide are an order of magnitude greater than the PHANToM. This makes it particularly suitable for tasks that require whole arm movements that may require larger forces than the PHANToM can supply. It is also a single point of contact device and has a larger workspace than the standard PHANToM models, but is less portable due to its size. Currently the end effector is pen-like, and does not measure rotation on any pen movements or provide force feedback in the rotational axes. It supports lateral motion, static contact, contour following, pressure and unsupported holding similarly to the PHANToM.

Figure 9. The HapticMASTER developed by FCS Robotics.

2.5.2.4 The Laparoscopic Impulse Engine

The Laparoscopic Impulse Engine [49] (shown in Figure 10) from Immersion is designed as an application specific device for simulation of minimally invasive surgery. It provides 5 input DOF and 3 degrees of active force feedback. The user can move the device about the simulated tool's insertion point, and can receive force feedback on these axes. Force feedback is also available when inserting the tool into the patient. The device allows rotation of the tool, and the opening and closing of the grippers although no force feedback is provided on these axes. This device has been specifically developed to simulate exploration of an object using a probe.

Figure 10. The Laparoscopic Impulse Engine available from Immersion.

2.5.2.5 The Cybergrasp

The Cybergrasp from Immersion [46] (shown in Figure 11) is an ungrounded force feedback device. It contains position sensors for the hand and fingers, and can supply 12 Newtons of continuous force to each of the fingers and the palm. Its design allows one degree of freedom for each finger. It therefore supports the EP pressure similarly to the Rutgers Hand Master II. Because of the multiple contact points, a limited form of enclosure is supported. However, lateral motion, unsupported holding and contour following are not possible with this device.

Figure 11. The Cybergrasp from Immersion.

2.5.2.6 The Wingman Force Feedback Mouse

The other devices described are designed for research and commercial use. The Wingman Force Feedback Mouse (shown in Figure 12) is an example of a product aimed at the desktop market. It is considerably cheaper but offers a lower quality of haptic feedback. It is similar to a standard mouse except it is constrained within a fixed workspace. The device provides two DOF force feedback. The forces supplied by the device are weak however and easily saturated. As it is a two DOF device, it allows for exploration of objects in two dimensions only. EPs supported by this device for two dimensional object exploration are contour following and pressure. Static contact is supported in a limited form only, as no temperature information is provided.

Figure 12. The Wingman Force Feedback Mouse.

2.5.2.7 SPIDAR

The SPIDAR force feedback device [10] is shown in Figure 13. The device provides force feedback to a user by mean of tensioned strings. Four tensioned strings are attached to a ring that the user can slip over his or her finger. Position encoders and motors attached to the end of these stings allow the device to track the hand position of a user, and apply force feedback. At the front of the device shown in Figure 13, a large screen is incorporated to allow a user to view his or her interactions. One advantage of this device is that the it can be scaled in size to suit the task at hand. If only arm sized movements are required, a smaller frame can be used. Figure 13 shows a large room scale SPIDAR. The user can interact with more than one interaction point at a time by adding another set of four strings attached to a ring. It is difficult to ensure however that the tensioned string from the different interaction point do not interfere with each other. The device can exert a maximum force of 30 Newtons, which is larger than the PHANToM.

The SPIDAR allows the user to perform similar EPs to the PHANToM and HapticMaster devices. Lateral motion, static contact, contour following, pressure and unsupported holding are provided by the device. Similarly to the PHANToM, static contact is only supported in a limited form as temperature information cannot be conveyed and the device lacks the tactile cues that are used during contour following. Enclosure is not supported. However, an extremely crude form of enclosure could be provided by allowing more interaction points.

Figure 13. The scalable SPIDAR force feedback device is shown on the left. The right picture shows a user interacting with the SPIDAR.

2.6 Current applications

There are many current research and commercial applications for haptic devices. Interaction with Virtual Reality scenes is one major area. Not only is touch feedback provided by haptic devices, but some devices also offer a higher number of degrees of freedom than traditional devices such as a mouse. The PHANToM for example can supply force feedback in three dimensions, but also allows intuitive navigation of a VR scene in three dimensions. One successful VR application is Freeform Modelling developed by SensAble Technologies. Freeform is a virtual sculpting tool for developing three dimensional virtual models. This system is employed by a number of companies throughout the world in particular to build product prototypes that can not only be seen, but touched also.

Haptic devices are commonly used in psychology research for investigating properties of our haptic system. These devices can offer accurate position and force measurements as well as controllable mathematical models of objects to enable researchers to gain insight into the haptic system. One example of such research is examining grasp forces, under different conditions. Pollick *et al.* [92] examine how the mass or friction with the environment of an object affects the grasp forces used to lift the object. Augrelle *et al.* [3] examine how different gravitational forces affect the grip forces when moving an object. One interesting area of psychological haptic research is in cross modal perception. It is rare in a real life scenario for a person to experience an object through one sense alone. It is therefore important to examine how the senses interact with each other. McGee [74] is examining the presentation of virtual textures through haptics and audio.

Haptics has been investigated for use as a training tool in many areas. Yokokohji *et al.* [121] have investigated the use of haptic guidance to provide training for simple assembly tasks. Sakuma *et al.* [98] have developed a system to train a user in calligraphy. The SPIDAR device is used to guide the user through the pen movements required to form Japanese symbols.

Haptic devices are also being researched as a tool for aiding visually impaired users. Braille displays have previously been discussed in this chapter, but researchers are examining the possibility of providing data such as statistical data to visually impaired people. Brewster [12] describes the Multivis project, which examines the use of virtual haptics combined with audio to present a visually impaired user with different forms of statistical information that would normally be presented in a visual form.

Tele-manipulation and remote communication is an area in which haptics can play an important role. During a tele-presence manipulation, the user is separated from the manipulated object, and the tool performing the manipulation. The operator may be receiving minimal information from the remote environment, which can make the task he or she is performing more difficult. Particularly when the user needs to know when he or she contacts an object, and for breakable objects, how much force is being exerted by the remote device. Haptics provides an extra channel of communication that can help the operator. Reinhart *et al.* [94] present a system where a user interacts with a PHANToM 6DOF device that manipulates a remote robot arm. The task is to place cogs into a clockwork mechanism, which requires great precision. The tele-manipulation allows large movements from the operator to be mapped to small movements from the robot arm increasing the margin of error. The haptic device can provide the operator with cues when the robot arm contacts other object in the clockwork mechanism. Haptics also has a role to play in remote communication. Oakley [87] has examined the use of haptics to increase presence and aid communication in a distributed collaborative environment. Similarly Sallnäs and Kjoberg [100] have investigated how haptics can improve task performance in a shared distributed environment.

One of the largest areas of haptic research is in providing systems for medical training. Chapter 3 will discuss the use of Virtual Reality in medicine, and in particular for medical training simulations.

2.7 Conclusions

This chapter has reviewed how humans use their haptic sense to explore and identify objects. Important definitions have been introduced, and experiments have been described that identify different conditions that will affect haptic object identification. The review of the literature presented in this chapter suggests that despite virtual touch being more limited than haptic exploration in the real world, it is possible to simulate virtual objects that can be recognised through interaction with haptic devices. The strengths and weaknesses or different forms of haptic devices have been examined and the conclusion can be reached that currently, force feedback devices are more suited than tactile devices to representing three dimensional virtual objects.

These factors are important when considering the first two research questions from Chapter 1. The limitation of various haptic devices must be taken into consideration when designing the simulation. Particularly for a palpation simulator, the device must allow the user to perform the necessary EPs. During a palpation examination, a doctor or veterinarian will be using his or her sense of touch to identify different properties of the contacted area. This will involve haptic identification of the contacted object's properties. It is encouraging to note that there are a wide variety of devices that support a number of he EPs described by Lederman and Klatzky [66]. The fact that virtual haptic object identification is possible, albeit in a limited form, is important when considering the second research question. A user of a simulator should be able to identify a simulated object with the real world object that is being modelled. The research described in this chapter suggests that this can be the case.

The following chapter will examine in detail the current uses of haptics in medicine. It will focus particularly on Virtual Reality simulators for training and assessment.

3 Virtual Reality in Medicine

This chapter contains a review of previous work relating to Virtual Reality in medicine. An overview is presented of some of the major uses with specific focus on Virtual Reality simulation as means of a teaching tool. This is relevant to all three key research questions presented in Chapter 1, as before working on these questions, it is necessary to take account of and build on previous work in the area. This chapter looks at the limitations of current training methods, and why the incorporation of medical simulators into training practices would bring benefits. It reviews current attempts at design, development, and validation of medical simulators. It argues that performance feedback is one major area of benefit that simulators can provide and reviews current attempts at providing performance feedback through Virtual Reality simulators.

3.1 Overview of uses

Virtual Reality systems are currently being researched for use in several areas of medicine, although there has so far been little widespread acceptance in the medical community. Originally, computer systems tended to concentrate solely on the use of the multimedia presentation of information. However, as computers have become more powerful, and new devices have been developed, Virtual Reality systems are now viewed as potentially beneficial in the medical domain.

Current research can be loosely categorised into five main areas: computer assisted surgery, telepresence surgery, medical visualisation, medical rehabilitation and medical training and education.

3.1.1 Computer Assisted Surgery

Virtual Reality for use in computer assisted surgery often takes the form of augmented reality systems. Augmented reality is a hybrid of digital and real environment spaces. In augmented medical procedures, additional feedback that is not normally available to the doctor is presented during the procedure to enhance the information the doctor receives. Edwards *et al.* [34] present an augmented reality system in which a surgeon can view preoperative radiological images accurately overlaid onto the patient in stereo through a surgical microscope. Using this microscope, a surgeon can plan an operation to avoid critical structures, or even locate tumours. Similarly, Liévin and Keeve [69] suggest a similar system to overlay images on a patient to allow the surgeon "X-ray" vision. The system is not fully implemented and uses a monitor to display the images, however, stereo vision glasses would eventually be used to supply visual feedback overlayed on the patient to the surgeon. Accurate alignment of the virtual images and real world objects has been identified as a critical feature of this system as alignment errors could mislead the surgeon, and degrade performance. If the overlayed image is displayed in the wrong place, surgical tools could be inserted into the patient in an inappropriate place, possibly harming the patient.

Wegner [120] demonstrates how audio feedback can be used to present state information in a medical context to augment a surgical procedure. The procedure used to demonstrate this concept is a neurosurgery procedure that requires the surgeon's visual focus to remain on the patient. This system presented surgical instrument position and optimal path information to the surgeon through audio, allowing the surgeons to use the information while keeping their visual focus on the patient.

Haptics can also play an important role in computer assisted surgery. Howe [45] notes the difficulties introduced when interacting with a patient through minimally invasive tools. In such an examination, a surgeon views his or her actions through a monitor, receiving only force feedback through the handles of a long tool. Sometimes referred to as keyhole surgery because of the incision made in the patient is small, the surgeon very much relies on visual feedback from the monitor and tactile feedback from the tool to guide his or her movements. By using a sensor at the tool tip, and a tactile display at the surgeon's end of the tool, it is possible to provide a tactile impression of the current object touched by the sensor. This extra sensory channel will provide more information to the surgeon allowing for better small scale shape perception. Kumar *et al.* [62] describe a system for aiding surgery requiring small movements. The paper describes how a robotic system attached to a surgeon's tools can sense and cancel out small scale tremors from the surgeon. This system has the potential to aid surgeons in making smoother small scale movements by reducing unintentional movements due to tremor.

There are potential legal issues that must be considered when assisting a surgeon. If a patient is harmed during a computer assisted surgery procedure, is it due to a mistake by the surgeon, or is it the fault of the tool that may have presented the surgeon with misleading information? This will be particularly important for force feedback systems that will directly affect the movement of a surgeon within the patient by either assisting or resisting movements in certain directions. Ethical issues make validation of these tools a difficult problem.

3.1.2 Telepresence Surgery

Telepresence surgery systems allow a surgeon to operate on a patient remotely. The surgeon would control a local device that would communicate with tele-manipulation equipment in the patient's locale [83]. Although no commercial systems are currently in use, many have been developed for research purposes. The Telepresence Surgery System TeSS [55] is an example of a system designed to operate on a patient over distance. The remote surgeon is provided with real time 3D video vision captured by two video cameras on the patient side and is presented with the information through a stereo display. The system was tested specifically for mentoring purposes, such that students could not only observe operations without the surgeon blocking their view, but could also feel the motions of the surgeon without affecting the movement of the of the surgical tools. The tests also showed that the remote surgeon was able to successfully complete all seven operations. Docimo *et al.* [31] demonstrated the effectiveness of a telepresence Minimally Invasive Surgery (MIS) procedure. The remote surgeon was provided with real time video and two-way audio communication and could manipulate the laparoscopic equipment by controlling a remote robot arm. Twenty-six of twenty-seven procedures were successfully completed without the aid of a surgeon in the operating room. Poor positioning of the equipment in one trial led to help being required from a surgeon in the operating room. It must however be noted that this procedure was performed over a short distance. Longer distances may introduce the problem of network lag. The large quantities of visual data sent, and the fast update rate required for haptic feedback will provide restrictions on these procedures over large distances with current network technology.

Telepresence surgery could prove to be useful in several areas. The American military view telepresence as a method of providing treatment to injured soldiers on a battlefield, [101]. A combination of sensors to monitor vital signs, and GPS devices can be used to transmit to some central medical station the condition and position of all soldiers. A surgeon will therefore practice triage through viewing the vital signs of the wounded soldier and select the most urgent case. If a soldier requires immediate surgery, he or she can be placed into a remote surgery station. The wounded soldier would then be operated on by a surgeon at a remote site without the risk to the surgeon or the risk of moving the patient. Similarly, patients living in remote areas could receive expert care from a central specialist without having to travel. Both of these situations would require specialist telepresence equipment to be present at the site in order to perform the procedures, but it offers a method to provide specialist care to a patient in areas where not normally possible.

There are however several concerns about telepresence procedures that have restricted its acceptance [31]. In the event of technical failure or complications a doctor would need to be present to complete the operation, which detracts from its usefulness. Malicious attacks on a network are also a problem that must be addressed. In surgical telepresence systems, security would be an essential component as any external interference could lead to injury or death for a patient. Security is also important for the purposes of patient confidentiality. Telepresence surgery requires that confidential patient information is transmitted over a network.

Licensing of surgeons may also present problems across country boundaries. Would the surgeon be required to hold a medical license to practice in a particular country before performing telepresence surgery across borders on a patient in that country? Or would it be sufficient to hold a licence valid in the country that the surgeon is in. This will require legislation before cross border telepresence operations become an accepted part of medical treatment.

Further questions need to be answered about its effectiveness. Will network lag over long distances affect the performance of the surgeon? This is particularly the case for procedures that require a great deal of accuracy and fine grain movement. Will performance be degraded by an impoverished environment? The surgeon will view the tool movements through a monitor as in minimally invasive surgery. However, his or her view of the patient may be restricted, and some spatial awareness of the location of the tool in the patient may be lost.

3.1.3 Medical Visualisation

Visualisation is a promising area for the use of Virtual Reality and in particular three-dimensional graphical modelling. Medical visualisation systems have been developed for training and education, as well as diagnosis. The 'Visible Human Project' [84] is the most prominent example of using visualisation for training purposes. For this project, a human cadaver was immersed in gel, frozen and sliced into thin sections. A three-dimensional computer model was then generated by scanning the sections into a computer, and assembling them. The resulting model allows a viewer to view all or selected parts of the human body in three dimensions. The models can also be exported to provide anatomically accurate models for use in medical training simulations [28]. It is important to notes that although the geometry of the model will be accurate, the haptic properties will need to be determined separately. This information will be accurate for the one body only. Different patients will have different characteristics. Factors such as weight can affect the shape of the human body. The 'Lucky the Virtual Dog' project [33] illustrates another method of building up anatomically accurate models. Image segments of a dog cadaver capture by CT and MRI scans were assembled and built into a three dimensional model of a dog. The purpose is to provide training for veterinary students in the spatial layout of the organs in a dog. The system allows a user to remove and replace organs to selectively view different regions. There is, however, no literature discussing the model's effectiveness or suggesting that this model has been used to teach. This technique of building three dimensional models has also been used by the Department of Radiation Oncology at the University of Southern California School of Medicine to build models of human anatomy.

The 'Glass Horse' [81] similarly presents users with a three dimensional virtual model. The horse has been designed using CAD tools, and animated to illustrate different conditions. For example colic - which may occur when the intestines of a horse become distended or twisted - can be animated by showing the movement of the intestine, and how it becomes trapped. This visualisation is not possible without such a tool as it is an internal condition, and is only viewed and treated by a veterinarian after it has happened. The 'Glass Horse' provides users a view of the spatial layout of the different organs within the horse, and can demonstrate various ailments. However, for each new condition demonstrated, a new animation must be built. A screenshot from the system is shown in Figure 14.

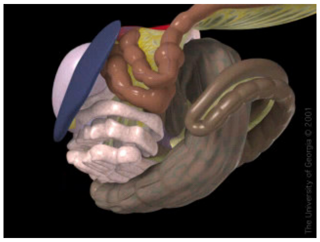


Figure 14. A screenshot from the 'Glass Horse' system. This image shows the organs present in a horse's abdomen and their relative positions.

One promising area of this research is to develop tools for importing models of a patient into the system before surgery. This would allow a surgeon to practice an operation with anatomically correct models before operating on the patient. It is important to note however that visualisation techniques would provide graphically accurate representations, but haptic properties would need to be provided through a generic tissue model. One example of such a system is presented by Keeve *et al.* [56]. This system provides a preoperative planning simulation for craniofacial surgery. Computer tomography is combined with a laser scan of the patient's skin to build up a 3D model of the patient's head. The surgeon can practice the procedure on the virtual patient model, and view predicted results of his or her actions before performing the procedure on the actual patient. The predictions are made through a mass spring model or finite element methods applied to a tissue model. A similar system has been developed by Radetzky *et al.* [93] for visualisation of brain tumours. Patient data is scanned in through standard imaging equipment for neurosurgical procedures. This data is then used to develop a volume rendered model of the patient's head. A surgeon can then use a specially built 4 degree-of-freedom device to practice the procedure on the simulated patient. As with all surgical simulations however, it is difficult to evaluate to determine if there is any benefit from it's use. Although this tool is potentially useful in planning and practicing nerosurgical procedures, there is no evidence to support the fact that it does provide benefit.

3.1.4 Rehabilitation

Medical rehabilitation is a promising area for Virtual Reality. VR has been successfully used to treat patients with phobias such as arachnophobia [20]. Patients are exposed to their fears gradually and systematically using a Virtual Reality headset to present the virtual world. As a patient's fears lessen, the virtual environment allows greater exposure of the feared object or situation.

Virtual Reality, and in particular haptics, is currently being investigated as a tool for aiding and assessing progress during physical rehabilitation. Injuries that restrict movement can require long term rehabilitation. Deutsch *et al.* [29] present the Rutgers Ankle System. This tool is a grounded six degree of freedom force feedback device that attaches to the base of a patient's foot. The strength of forces supplied by the device can be adjusted to suit the needs of the patient. A case study was presented that examined improvement of a stroke patient using the Rutgers Ankle. The patient controlled a plane, with the task set to fly through a series of hoops using the haptic interface as an input device. Improvements were shown in the mobility of the patient, and range of movement. Future plans exist to extend the system to send data from a remote system such that a physiotherapist could examine patient data and adjust the training schedule and forces used remotely. The VR system provides entertainment during the training sessions through the task set. This game is used to maintain the interest of the patient. Using a VR system for rehabilitation allows very accurate measurements of performance, along with a history of past performance such that trends in improvement or otherwise can be noted and the rehabilitation schedule adjusted accordingly.

Loureiro *et al.* [71] present a system that uses the HapticMASTER to provide force feedback for rehabilitation of stroke patients. Similarly to Deutsch, they identified motivation as one of the key aspects of the rehabilitation process. The task set to a user involved moving virtual objects around a virtual environment. The study used three different virtual rooms that were presented to the participants using visual and haptic modes of interaction. A questionnaire was used to study and assess the participants' perception of the system. The responses suggested that during the study, all 6 participants were motivated to use the system. Of the three rooms presented, participants preferred the room that was based on the real world environment around the patient. This is used to suggest that augmented reality is one area that should be explored as an area of research that is useful for rehabilitation applications.

3.2 An Overview of VR Simulators in Medical Training

Medical simulator training is a rapidly growing area of research. As this is the main component relevant for this thesis, the current literature will be reviewed in more depth in the following section. It will first give a brief overview of the use of simulators in areas other than medicine. Aviation will be used as an example of simulators that have been accepted as a valuable training tool.

It will then provide a comprehensive review of research and commercial systems that are currently available and will describe the different challenges involved when simulating minimally invasive surgery, surgery or palpation procedures.

3.2.1 Rationale

The use of VR medical simulations for training purposes is a rapidly growing research area, and as computing power and simulator fidelity increases, simulators are expected to become more widely accepted as a training aid. Flight simulations are often used as an analogy in that they provide training in a multi-dimensional, safety-critical task. Virtual Reality simulations have been used successfully for several years in flight training. Although not initially widely accepted, improved technology for flight simulators has lead to more realistic simulations that have proved useful to developing, maintaining and assessing pilot skill [96]. They have been successfully employed to simulate a wide range of conditions and emergencies, while reducing the learning curve for trainee pilots by providing a safe environment in which to learn. Since their introduction, they have become an integral part of pilot training. Simulators have also proved useful in providing training for situations that are prohibitively expensive to train for otherwise. For example, NASA is committed to a 'total simulation' paradigm for its astronauts, in which the objective has been to train astronauts by simulating as much of the mission as possible.

Simulation training is not a new idea in human and veterinary medicine. Students gain experience in certain techniques through use of plastic or rubber models, but these can be expensive and often lack realism while providing no useful feedback to the trainee. Barker [5] observes that students often resort to practicing venipuncture (inserting a needle into a vein) on each other, as the soft plastic moulded arms used for training do not provide sufficient realism. Although realistic models may exist for some of body parts such as bone, surrounding tissue is often unrealistic and lacks physiological properties of real tissue [44]. Physical models are often static and often model only a small part of the task. Surgical skills can also be improved in the anatomy labs that are incorporated into the medicine and veterinary medicine courses. Again, there are problems however since cadavers are a scarce resource, and are not generally reusable. Living tissue may also have noticeably different haptic properties than cadaver tissue [44]. There is also the problem that cadaver models teach the anatomy but not the physiology of the body. Advanced Medical Simulations [1] is a company that specialises in providing advanced physical models with computer enhancements for training purposes. A user can practice skills such as ultrasound scanning on a mannequin, with the appropriate feedback data being shown on a monitor.

Virtual Reality simulation training has many potential benefits in the medical field. The models used in a simulation are reusable unlike cadaver models, so after the expensive initial equipment cost of the simulator, costs would be reduced as there is no need to replace models. Most importantly simulators provide students and experienced doctors alike with a safe, controllable environment in which to learn new skills as well as practice familiar techniques. Traditionally, medical training has used the apprenticeship paradigm, where a student observes several operations performed by an experienced doctor before trying the operation himself or herself under the supervision of an experienced doctor. There are obvious risks to the patient, as a student may not have the necessary psychomotor or cognitive decision making skills necessary to perform the operation. Medical simulation offers the possibility of providing a method of learning a procedure by practicing it many times without putting a patient at risk [43].

Difficulties exist when trying to provide training for rare conditions. For both medical and veterinary medical training, students can only be exposed to the conditions that occur during their training. When a medical practitioner becomes qualified however, he or she would be expected to diagnose and treat conditions when all previous exposure may have been through lectures or reading text books. Similar problems exist with invasive procedures such as prostate examination. Prostate examinations can be very distressing for a patient, and ethical issues prevent allowing several students to examine one particular patient. A virtual simulation for a rare or invasive procedure would allow larger access to inexperienced clinicians for these procedures where they traditionally get little practical experience.

3.2.2 Current Systems

In Section 3.1.3, the 'Glass Horse' and 'Lucky the Virtual Dog' were introduced as visualisation tools to be integrated into veterinary course. These examples are designed for visualisation. Not to allow students to practice their skills. There are no examples of medical simulators being designed specifically for veterinary medical training. This review will therefore concentrate on existing systems available in human medicine.

The current virtual medical simulations can be separated in to three separate disciplines: minimally invasive surgical simulations, surgical simulations and palpation simulations. While the goals of these systems maybe similar, each of the types of simulator present different challenges to developers, and as such will be discussed separately.

3.2.2.1 Minimally Invasive Surgical Simulations

Minimally Invasive Surgery (MIS) procedures are by far the most common procedures simulated [14, 61, 73, 105]. In a MIS procedure, surgeons view their interaction with the patient through a monitor, and hence it lends itself to a virtual simulation viewed through a computer screen. MIS surgery is often referred to as keyhole surgery as the tools used enter the body through one of more small incisions. These procedures are becoming common as they offer the potential of causing less trauma to the patient, and therefore shorter hospital stay times. MIS procedures however require a higher level of dexterity from a surgeon than traditional surgical procedures [25]. This is due to the fact that there is a mapping that must be learned between the surgeon's movements and the movements at the tool tip. There is a fulcrum effect caused by operating through a small incision such that a movement from the surgeon from right to left will correspond to a movement of the tool tip from left to right. There is evidence to suggest that some formal training mechanism is required, as Royston [97] notes that in 1994, around one third of all MIS surgeons were self taught. The most commonly modelled procedures are endoscopic, laparoscopic, or arthroscopic. These different types of MIS procedures may have different requirements for simulation depending on the area of the body being operated on.

Endoscopy is an exploratory procedure performed with an endoscope - a minimally invasive tool with a camera at it's tip. One example of an endoscopic simulator is the Preop endoscopic simulator [14] developed by HT Medical Systems. It is a system that combines force feedback through a passive force feedback device - to minimise cost - with anatomical and physiological models of the human bronchial tubes. This is a procedure where the physiology is an important aspect of the examination. For example, the patient will cough if the part of the bronchial wall being touched is not anaesthetised properly. The user can then choose to anaesthetise that area. Mucus from the bronchial wall will obscure the surgeons view as in the real procedure. These are features that would be difficult to implement in a physical model. This system not only simulates the endoscopic procedure, but provides guidance for pre and post-operative skills as well. This recognises the fact that a successful surgical procedure relies on successful completion of pre and post-operative tasks. Preop is a commercially available training system, although no evidence is presented in the literature to suggest it provides training benefits to users in endoscopic procedures.

Laparoscopy is minimally invasive procedure performed on the abdomen. This area of the body contains mostly soft tissue that will deform when pushed or pulled by a laparoscope. Tendick *et al.* [113] present a system to provide training for laparoscopic surgery. Their system incorporates modelling of rigid and deformable models. They have identified interaction with the tools as a key component of the training, and as such have developed a system that trains users to move the tool tip to different points in a virtual environment. In this instance, the environments are not necessarily medical, but will require the user to move to different target positions.

The Minimally Invasive Surgery Trainer (MIST-VR) available from Mentice is one of the most established commercially available Laproscopy training systems. Cosman *et al.* [22] provide an overview of this simulator. It is designed to train surgeons in the core skills involved in minimally invasive surgery such as movement of the tools and suturing, and is primarily aimed at improving psychomotor skills required for performing this type of surgery. The simulated tasks are abstracted from the real procedure. That is, it consists of 12 individual tasks that individually require similar movements and techniques to the real procedure to complete, without necessarily looking like the real life procedure. These tasks include such core skills as tool manipulation, suturing, and diathermy (the generation of heat in the tissue through an electric current for surgical purposes). The system provides graphical feedback to the user but no haptic feedback.

Arthroscopy is a minimally invasive procedure performed on joints. Arthroscopic simulators must therefore model rigid objects such as bone. Similarly to laparoscopy, arthroscopy simulators must be able to handle cutting and suturing of structures. One example of an arthoscopy simulator is the Sheffield Knee Arthroscopy Simulator (SKATS) presented by Arthur *et al.* [2]. SKATS was designed with the aid of surgeons to train users in knee arthroscopy. The simulation consists of a high-resolution visual representation of a knee joint. A user is placed in realistic conditions and interacts with the virtual patient through a modelled arthroscope. However, testing of the system highlight problems as users concluded that haptic feedback was important in completing the task. Surgeons often use haptic cues to navigate the probe through the body in MIS procedures, and the fidelity of the simulation suffered with the lack of these cues. Other systems exist to simulate arthroscopy [105].

3.2.2.2 Surgical Simulations

Surgical simulations cover a wide range of techniques using different surgical instruments [5, 6, 9, 28]. The requirements for surgical simulation involving different sets of instruments may be very different. Cathsim [5] is an example of a commercially available training system for venipuncture. Venipuncture is the process of puncturing a vein most often for he purpose of collecting blood or administering medication intravenously. The system models insertion of a needle in to a simulated vein. Initial tests show Cathsim to be a popular alternative to the current rubber arm models provided for training, although no direct evidence is provided in the literature of any training benefits. Berkley *et al.* [9] present a simulation for training in wound suturing. Suturing is an essential skill for a surgeon to learn. The models involved in simulating a suture will be complex as the tissue needs to be both moveable to close the wound and deformable such that the skin will stretch realistically when pulled. Berkley therefore chose a finite element method approach to model the tissue. Finite element methods (described in more detail in Section 3.3) are highly computationally expensive, but are considered the most accurate means of modelling soft tissue.

Simulation for cutting procedures in particular present different challenges, as models need to be dynamically adjustable to allow incisions. The surgeon may be required to cut open a virtual object, or even remove a section of tissue. Delp *et al.* [28] describe the development of a simulator to simulate an operation to repair a gunshot wound to the thigh. The thigh model used has been imported from the Visible Human Project [84] so is anatomically accurate. The simulation provides functionality for different surgical instruments. A scalpel is used to make incisions or to sever damaged tissue. Forceps and haemostats - used for grasping objects during surgery and clamping blood vessels to stop blood flow respectively - are provided to grasp fragments of shrapnel and severed tissue. A wound probe to examine the trajectory of the bullet is also provided. The surgeon can select the instrument from an instrument panel in the corner of the screen. Delp *et al.* also present cutting and bleeding models they use to provide the simulation with the appropriate functionality. Cutting is a feature of the scalpel tool and is performed by boolean subtraction of the scalpel area from the currently contacted object. A separate object is generated when the scalpel completely severs the tissue. The bleeding model is based on fluid dynamics, and can model the oozing of blood as well as pooling in sunken areas.

Satava and Jones [102] are in the process of developing an ambitious system for simulating all aspects of battlefield medicine. This is to be incorporated into a larger battlefield simulation. The medical doctor would be part of a platoon of soldiers. They view the environment using a head mounted display, and explore using a stationary bicycle for movement. When casualties are detected, the doctor must practice triage to select the cases requiring the most urgent attention, and then perform the surgical procedure required on the soldier. This system introduces several technical challenges in representing interaction with the environment in a natural manner, particularly for transition between the different tasks (such as movement and examining a patient). Also, a wide variety of gunshot wounds must be modelled as well as provision for a large number of necessary procedures. The total simulation paradigm is an important issue, as Satava and Jones note that the battlefield doctor requires more than just surgical skills.

3.2.2.3 Palpation Simulations

The development of a palpation simulation presents different problems than a surgery or MIS simulation. During surgery, a medical practitioner interacts with the patient through surgical instruments, so the instruments mediate the haptic feedback from the tissue to the surgeon. Palpation, however, involves the medical personnel interacting directly with skin or tissue. The development of palpation simulators is less common, although palpation is an important technique for early diagnosis of many conditions. Recent examples have been developed at the Human Machine Interface Laboratory at the CAIP Center at Rutgers University. Dinsmore *et al.* [30] have developed a simulation using the Rutgers Master II [18] for training in palpation for the detection of sub-surface tumours using experimentally based force-deflection curves. Sub-surface tumours in their early stages are often detected through palpation. The system models tissue using experimentally generated force deflection curves to represent object rigidity, with tumours being modelled as more rigid than the softer skin. The region of palpation was restricted to the abdominal region. The skin can be set transparent such that the user can observe the organs inside while palpating.

Burdea *et al.* [16] present a palpation simulator, for modelling the human prostate. The system renders objects graphically and haptically, with the force feedback being provided by the PHANToM. The simulator can model several different prostate tumour conditions from single tumours of different sizes to clusters.

Stalfors *et al.* [110] present a system where surgeons can palpate subsurface tumours. This system is intended to allow surgeons to remotely palpate scanned in tissue data from a patient in order to decide whether to operate to remove the tumour or not. One thing not address in this paper however is the importance of palpation if the data is actual patient data. As the data has been scanned in and a virtual model created, it would seem to open up the possibility of using the system as a diagnostic tool, and thus removing the need for palpation by a surgeon. They do, however, note the possibilities of using the system as a training tool for novices. There is no data presented to support the effectiveness of the system as a training tool.

3.3 Simulator Fidelity

Satava [102] notes that realism in simulation is the prime determinant of believability. The fact that users are asked to suspend their disbelief attests to the immaturity of the technology and the need to improve. Arthur *et al.* [2] provide four factors influencing the fidelity of a system:

- Physical fidelity is the degree to which a simulation looks like the simulated object and its control inputs have the same behaviour.
- Operational fidelity is the degree to which a simulator will actually operate in the same way as the simulated object.
- Functional fidelity is the degree to which it is possible to carry out the same tasks in the simulation as in the real task.
- Motivational fidelity is the degree to which the simulator is acceptable to users and valued as a source of legitimate and validated training. This will be affected by the above fidelity factors, although will also depend on such things as usability of the system, accessibility, and cost.

A system has a finite amount of computational power, and these factors can conflict. Choices must be made as to how to provide the greatest fidelity. For example, greater resolution or realism of the virtual models used might be balanced against the introduction of another modality into the simulation. The SKATS [2] simulator is an example of such a situation. The aim was to create knee arthroscopy simulation to provide training to novice and experienced surgeons. The choice was made at the requirements capture stage of the project to provide visual feedback only, which eventually limited the use of the simulator for training purposes. During the user testing stage, surgeons identified the need for haptic feedback as in arthroscopy, haptic cues are often used to explore.

Simulators also have the potential to present anatomical and physiological information to the user simultaneously, but the physiological information is often ignored as it can be complex to model. Fidelity will also be affected by the choice of method for object modelling. Anatomically accurate data will provide greater fidelity than approximations to the model. Choice of algorithm for haptically rendering of soft objects will affect fidelity also. Briggs [13] suggests that current linear tissue models may mislead a user and provide negative training. But the question remains of how to measure the haptic properties of live tissue.

There is a large area of research into increasing realism in simulators by improving soft tissue properties. A large body of work relates to modelling objects that allow more realistic haptic and visual deformations. In simulating a deformable model, the system must calculate the forces generated against the user's interaction points (the object's reaction force) as well as the modified shape of the deformed object.

There has been considerable demand for research into the modelling of soft or deformable objects, particularly from the medical simulator community. Tissue models have become more realistic as computational power has increased, but there still exists a trade off between complexity of the deformable model and computational power required. High servo loop rates required for haptic interaction limit the computational complexities of these models. More realistic tissue models also tend to be more computationally expensive. Modelling tissue presents problems as it is non-linearly deformable, and may also exhibit anistropic behaviour [13]. That is, the tissue may exhibit different properties depending on the direction of force. Cugini [23] observes two common approaches to modelling deformation: geometric and physically based models.

3.3.1.1 Geometric-Based Models

Geometric models focus on the appearance of the object, but ignore the physical properties of the tissue. They tend to have low computational requirements, but are not based on physically accurate principles [23]. The range of deformation from the contact point is chosen arbitrarily, and the reaction force from the object is calculated by a predetermined formula [28]. The simplest and most commonly used method for rendering reaction forces for compliant object is the Hookes-law model.

Reaction Force = Spring constant * Penetration Distance

The reaction force from an object varies linearly as the penetration distance of the interaction point into the object. A compliant object can be modelled by a low spring constant. This can be used to model spring behaviour and requires little computation, but is a simplistic model for tissue. Burdea *et al.* [16] introduce a method based on experimentally calculated force-deflection curves that has been incorporated into their palpation training simulator. They measured experimentally the deformation of a rubber prostate model with different pressure forces applied and generated a quadratic equation to calculate the reaction force for different penetration distances of the interaction point within the object. The result was a fast non-linear geometric model, based on physical calculation. However, one failing of Burdea's method was that the calculations were based on a rubber prostate model rather than an actual prostate. Obvious ethical issues prevent the same method being used on organs in living person. Geometric models are considered not adequate to described complex non-rigid objects. This method also ignores the anistropic properties of tissue. For these reason, current research focuses on physically based approaches.

3.3.1.2 Physically-Based Models

"A physically-based model is a mathematical representation of an object, and of its behaviour that incorporates forces torques, energies and other attributes of Newtonian physics" [23]

In physically based models the geometry becomes time and force dependent. These models can be used to describe not only the geometry on an object, but also physical properties such as elasticity and viscosity. Different approaches have been taken to physical soft body modelling.

Finite Element Methods (FEM) is generally accepted to be the most promising approach to model tissue [9] [4]. FEM is commonly used in mechanical engineering for analysing structures and common applications including static, dynamic and thermal behaviour of physical systems. This method considers a deformable object to consist of a collection of finite-sized particles. The particle behaviour and overall structure is obtained by formulating a system of algebraic equations that can be solved by a computer. However, FEM is a very computationally expensive method and is not generally used in situations requiring real time rendering. Berkley *et al.* [9] demonstrate that FEM deformations can be achieved in real time by limiting the contact scenarios. Their method assumes only one contact point exists between the user and the object, and that this contact point exists on the object's surface.

De Angelis *et al.* [26] suggest an approach where objects do not exist as primitives, but as a series of particles distributed in space related by internal forces and constraints. This approach provides more flexibility than the primitive object approach as non-uniformly dense objects can be modelled by redistributing the particles. Forces among the particles are modelled as a set of equations that can integrated through time to determine the geometry. De Angelis notes however that this method is for now prohibitively computationally expensive for all but the simplest cases involving few particles.

Physically based models offer a method of visually and haptically rendering deformable objects more realistically than geometry based models. However, it is generally much more computationally expensive. A novel approach is taken by De [27] to reduce computation by modelling an object as a "thin walled membrane". De observes that during object deformation, we see the surface deforming and feel the reaction force due to the pressure applied on the object. By reflecting the properties of the material components inside the object, De suggests it can be reduced to a surface model. The method used represents the inside of the object as a fluid. De observes that efficient models exist for solving the behaviour of these thin walled structures. This method can be used to represent a wide range of objects even in the medical domain. Compliant organs like the stomach or spleen can be modelled in this way.

All these methods handle deformation of a single object. Kuroda *et al.* [63] note that this is a simplistic view in most surgical instances. The human body contains many deformable organs that will interact with other organs around them. This will again lead to added complexity when attempting to realistically model deformations. Further computational power will be required if cutting and bleeding is introduced is required to model the procedure. With current computational power, approximations and assumptions must be made when modelling tissue.

3.4 Validation of a Simulator

It was stated in Section 3.2.1 that simulators offer the potential of providing a risk free environment for clinical practioners to train without damaging a patient's health. It can however be argued that the risk still exists, and is merely shifted onto the design and validation processes involved in developing the simulator. If a simulator fails to provide the training it was designed to provide, inadequately trained practitioners may still perform procedures on patients. In the worst case mis-training could occur, where a simulator trains a user in a method that will degrade his or her performance in the real procedure. Validation of a medical simulator before use as a training aid is therefore essential.

The aim of a medical training simulation is to provide cognitive and psychomotor skill transfer to the user that is appropriate to the modelled procedure [121]. A badly designed simulator may provide no benefit or in some cases provide negative skill transfer, adversely affecting the user's performance in the real world task. A practitioner may be given a false sense of confidence through using a poorly designed simulator about his or her ability in performing a procedure.

There have been several attempts to validate the use of simulators with varying success. Higgins *et al.* [44] outline a methodology for the design and validation and eventual certification of a medical training simulator shown in Figure 15. There are two main sections described in this method: design and validation.

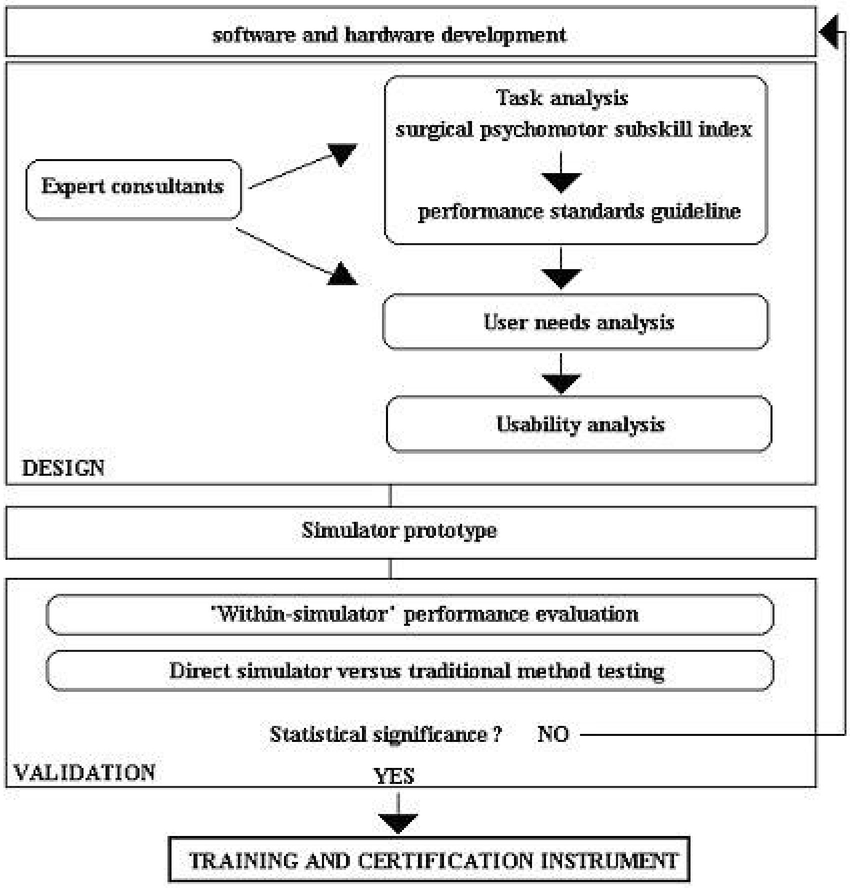


Figure 15. This outline shows a process by which a medical simulator can be validated as a training and certification instrument [44]

The design phase of the project is split into subsections. A standard task analysis is used to identify the different steps involved in the surgical procedure. The analysis is used to identify:

- The tasks that constitute the sub-steps of a specific feature (such as the surgical operation)
- The steps that contain critical cognitive and motor training. These are features that must be included in the simulation

The user needs analysis identifies the optimal feature set that the user requires to complete the specified tasks. This can be achieved by observation of a procedure, or interviewing an experienced surgeon. With the usability analysis, the designer aims to ensure that the simulation can be used as required. The interface to the simulator must not affect the user's performance of the task. In extreme cases, irritating features of the interface may discourage use of the simulator altogether. Higgins emphasises the importance of having input from experienced surgeons at the design stage. This is not only to guarantee realism, but also to provide face validity for the larger community of surgeons.

Validation of the simulators does present problems. Before the simulator is validated, it is not known if it will provide any form of training in the procedure. There are therefore ethical issues in directly comparing the performance of students trained with traditional methods and students trained on simulators. The method used by Higgins is with-in simulator performance, where experienced surgeons perform the procedure using the simulator, with their simulator performance being compared to their real world performance. Other attempts have been made to validate medical simulators although there are no examples in the literature of a medical simulator being developed using this methodology.

Validation of simulators will eventually lead to certification of these devices. The Medical Devices Agency, who ensure medical devices in the UK comply with strict European Union rules has stated that no certification of simulators exists as yet, but this would seem to be an important factor before they can be more widely accepted as a training aid. Simulators are currently viewed similarly to other training aids, such as courses and books. These may be reviewed in professional journals, or endorsed by recognised experts, but are not generally certified or approved.

Neufeld and Norman [85] describe several measurable factors to assess a clinical educational or assessment tool that must be taken into account if the tool is to perform the task required. Their definitions relate to measurable factors that can be used to show clinical competence. He states that the various factors that should be tested are credibility, comprehensiveness, precision, validity and feasibility.

- Credibility or face validity describes the extent to which the instrument seems to test what is being measured.
- Comprehensiveness or content validity describes the extent to which an instrument samples the area of competence under consideration.
- Precision or reliability measures the extent to which the instrument returns similar results under different conditions, such as repeated measures or different users.

Validity testing shows the extent to which the test measures the factor that was intended to be measured. Face validity and content validity have already been covered, but Neufeld notes that other forms of validity exist.

- Construct validity examines performance of the tool in the hands of various experience levels of clinician. Good construct validity would be demonstrated by experienced clinicians performing better on the simulator than novice clinicians.
- Concurrent validity involves the comparison of the factor being measured with the best available external measure. The measures should happen concurrently.

- Predictive validity similarly involves comparison of the measured factor with some external measure, but this external measure is taken at some time in the future.
- Feasibility is an important and often ignored feature in training tools. Feasibility is a combination of affordability of the tool, and the ability of an institution to implement the tool in terms of scheduling and logistics.

Neufeld and Norman further note that on top of these factors, educational considerations need to be taken into consideration. These educational considerations can be applied to the use of simulators in a training course. Firstly, a simulator must be used in an appropriate way. That is, it should only be used in a course to train for procedures - or parts of procedures - where it has been shown to provide benefit. The user must be aware of the purpose of the simulator. They should be made aware of it's strengths and shortcomings. Finally, care must be taken with the side effects of introducing a system into the course. It will either result in more work for the student, possibly to the detriment of other areas of the course, or will need to replace one area of the training removing that area from the course.

Berg *et al.* [7] use Neufeld definitions in order to discuss issues involving validation of a dermatologic simulator. They state a simulator should have face validity for the simulated task. It must look and feel like the objects being modelled. This is important in order for clinicians to accept the use of the training simulator for the task. Reliability is also important for a training simulator. A simulator should return similar results when tested over repeated sessions or with different skill levels of user. Content validity for a simulator would measure the extent to which the simulator samples all possible aspects of the skills in the area being modelled. A simulator with low content validity would only simulate a small section of the modelled task. A simulator with concurrent validity would need to demonstrate a correlation in a user's performance on a simulator with respect to some external measure. For example, skill level in the real life procedure could be compared with skill level on the simulator. Good predictive validity, would be shown for a simulator if performance on the simulator correlated with future performance in the real task modelled. Predictive validity - unlike concurrent validity - cannot therefore be measured at the time of training. Construct validity examines performance on the simulator between clinicians at different levels of training. Good construct validity would be demonstrated by experienced clinicians performing better on the simulator than novice clinicians. For a simulator to be feasible, it must be affordable, and should be able to integrate with the current medical course.

Gorman *et al.* [42] present a surgical simulation for driving a simulated needle through a target overlaying a blood vessel. To quantify the performance of a user, they defined the metrics: time on task, accuracy, peak force applied, tissue damage, surface damage, and angle error. However, they note the difficulty in calculating tissue or surface damage accurately. Initially, eleven surgical residents with no prior experience performed a standardised training program using the simulator for fifteen minutes a day for three days. A significant performance increase was noted, based on the defined metrics. A second test involved five surgical novices practicing twenty minutes a day. Again significant increase in performance was noted. The subjects eventually reached a skill plateau, where performance did not increase further. This shows that psychomotor learning took place, although there was no evidence to suggest that the skills learned carried over to the real world task.

Smith *et al.* [109] attempt to validate the Prosolvia shoulder arthroscopy simulator by comparing the performance of orthopaedic surgeons, other surgeons, and medical students. The task involved manipulating the arthroscope to touch targets placed within the virtual joint. The metrics defined were: time to complete task, number of collisions between the probe and tissue, number of dangerous collisions, and path length ratio compared with the ideal path. The results showed that the orthopaedic surgeons were significantly faster than the other groups. However, this is obviously not the important measure when performing surgery. The orthopaedic surgeons also registered more collisions and a worse path length ratio. Closer inspection of the procedure revealed that contact with objects to test the integrity of the object is a common technique in arthroscopy. The system designers were not aware of the performance criteria beforehand. This emphasises the need for careful selection of metrics for any given procedure.

The MIST-VR simulator is one example of a simulator that has been evaluated more than once in the literature. A recent study conducted by Kothari *et al.* [60] compares the performance of this simulator against the Yale Laparoscopic Skills Course - an accepted medical training course. This course consisted of 3 tasks being performed on a physical simulation once a day over a five day period. The tasks performed were:

- Rope pass drill - This task consisted of using the laparoscopic trainer to grasp a rope banded with several colours. Laparoscopic tools were controlled by both arms, and the goal was to grasp each of the banded colours alternately with the tools in the left and right hands.
- Cut drop drill - This involved grasping ten bean shaped objects with the tool in the non-dominant hand and then dropping the object into a hollow cylinder.
- Triangle transfer drill - this task consisted of 5 triangles mounted with a loop. The task required grasping the loop and moving the triangle to a different location.

The MIST-VR training involved the user performing six tasks each day over a period of five days. These tasks were

- Acquire place - This task required the user to pick up a sphere and place it within a certain area.
- Transfer and place - This task required the user to pick up a sphere, transfer it to the tool controlled by the other hand and then place it within a given area.

- Traversal - Similar to the rope pass drill, the user was required to grasp a target and traverse it using hand over hand transfer.
- Withdraw and insert - The user was required to withdraw a tool from an operating volume, and replace it accurately.
- Diathermy - Diathermy is generating heat in tissue by electric currents. This task required the user placing the tool in the correct location and only applying the diathermy when the position was correct.
- Manipulation and Diathermy - This task required several of the skills described above. The user had to acquire a target and then apply diathermy to the correct place on the surface.

Before taking part in the training, all 29 participants were taught how to perform an intracorporeal knot on a piece of foam. This involved passing a needle through silk and a piece of foam, and then tying a surgeons knot. They performed six of these procedures with the time taken being noted. Eleven participants completed the Yale Laparoscopic Skills Course, and 13 participants completed the MIST-VR training. Both groups then tied six intracorporeal knots in a piece of foam. Results showed that both groups performed the task in a significantly faster time, and there were no significant time differences were noted between the two groups. In this study, no other information was noted about the intracorporeal knot tying skills of the participants other than time. A reduction in time may show an improvement in the motor skills required to move the laparoscopic tools. However, it is not the best metric when evaluating the success of a procedure. Other such factors such as knot strength or tissue damage should have a far greater effect on the perceived competence of a practitioner. One other factor that must be taken into consideration is that this study cannot differentiate between improvements due to performing the initial knot tying procedure six times, and the improvements due to training using the different techniques. This practice effect could have led to the time improvement noted in this study.

Gallagher and Satava [38] present a study testing the construct validity of the MIST-VR surgical trainer. A group of twelve experienced surgeons - who had completed at least 50 minimally invasive procedures - and a group of twelve novice surgeons were asked to carry out the same ten tasks. The tasks were tool manipulation tasks using the impulse engine to interact with the MIST simulator, and required similar skills in manipulating the tools to those required to complete a minimally invasive procedure. The results show that the experienced surgeons performed the task significantly faster than the novice group. The experts recorded a mean of fewer errors, and on average took shorter path length than the novice group. This suggests that the MIST simulator could be used to determine minimally invasive surgical skill. However, a study by Paisley *et al.* [91] provides somewhat conflicting results. Their study could not demonstrate construct validity for the MIST-VR simulator. This study is one of the most comprehensive attempts at validation of surgical simulators. It examines the construct and concurrent validity of six MIS simulations (five physical simulators and MIST-VR). Twenty six basic surgical trainees and sixteen experienced consultants took part in the study. The six tasks used in the study were:

- Knot security - This involved using MIS tools to tie a knot around a cylinder of 5cm diameter using silk thread. Knot strength could be tested using a tensiometer.
- Targeted Suture Placement - This involved placing a silk suture through 10 marks on material held taut. Deviation from the marked targets could then be measured.
- Skin laceration suture - Participants were asked to suture a laceration in a simulated skin pad using MIS tools. Performance could be assessed using a task specific checklist.
- Small intestinal anastomosis - Participants were asked to join up a severed artificial intestine using MIS tools. This task involved suturing and knot tying skills and performance could be assessed through a task specific checklist.
- Laparoscopic box trainer - This task involved picking up a polo mint (manufactured by Nestlé) in a position indicated by a coloured band, switching it to the tool in the other hand and placing it on a marker pin. Participants had to transfer and place five mints. Errors were noted if the mint was dropped, or picked up by any part other than indicated by the coloured band.
- MIST-VR - The transfer and place task and diathermy tasks described above were used for this study.

Both groups performed all the tasks. Of the six tasks performed on the simulators, no significant time differences noted between expert and novice groups. The Laparoscopic box trainer section of the simulator demonstrated a significant reduction in errors for the expert group, however, no significant performance differences were noted in any of the other five tasks. A weak correlation was shown in the novice group of duration of basic surgical training and time taken to complete the small intestinal anastomosis and MIST-VR tasks. The novice group performed the simulator tasks again six months after the construct validity experiment. The suture placement task was performed significantly more accurately than in the initial study. The intestinal anastomosis and box trainer tasks performed significantly faster. The time taken to complete the MIST-VR simulator task was significantly less in the second trial, and the movements of the participants were significantly more economic. There was no control group in this section of the experiment, so one factor that cannot be measured is the effect that the initial trial had on the performance in the second trial. Improvements noted may have been due to a practice effect of performing the task before albeit with a long time gap in between.

Langrana *et al.* [64] attempted to provide a validation for a tumour palpation system presented by Dinsmore *et al.* [30]. They examined the performance when using the system of 32 participants who had no medical experience. A group of 16 participants were allowed to feel the hardness of two types of tumour for 1.5minutes. The second group of 16 participants were allowed the same training, but for 5 minutes. When then presented with 6 cases with one or no tumours, no significant performance or time differences were detected between the two groups. Both groups however were able to locate over 90% of the tumours. There is no data presented to link these results to real world performance.

Burdea *et al.* [16] describe their attempt to validate a prostate simulator. The experiment consisted of three groups: Non-medical students with no prior prostate knowledge, urology residents, and a control group of urology residents. The non-medical students were introduced to the basics of prostate training through rubber models. They then practiced on the simulator with different tumour conditions before the experiment. The first group of urologists also practiced on the simulator for five minutes before starting the experiment. Each subject in the first two groups was then presented with twelve cases modelling a prostate with or without tumours and asked to diagnose them. The control group carried out the same procedure using rubber models instead of the virtual models. The results showed that the non-medical students performed slightly better in diagnosing prostate tumours than the urologists on the virtual models. The control group of urologists performed significantly better using the rubber models than the other groups. Although not stated in the paper, it is likely that as the control group was composed of urology residents, they had previous experience of training on the rubber prostate models. If this was the case, the control group could reasonably be expected to perform better on models that were familiar. Burdea *et al.* conclude that virtual models may still prove useful to provide training for prostate examinations, but the current models used are not realistic enough.

3.5 Performance feedback

The apprenticeship model in current use very much relies on the opinion of the experienced clinician as to how far along his or her training the novice clinician is. The method used by the novice for the procedure will reflect the method preferred by the experienced clinician, and this will not necessarily be the optimal method. This system also relies on subjective assessment of the performance level of the trainee by the supervisor. Virtual Reality simulators are now widely thought to offer the potential of providing a new medical training paradigm. As such, commercial as well as research systems are being developed worldwide. One of the major considerations in building a training simulator is how to provide performance feedback to the user. Higgins *et al.* [44] state:

> *"it is pointless to build a training simulator that doesn't provide useful feedback on performance to the trainee"*

When medical personnel are learning a new procedure, they must be aware when a mistake has been made so that they may adjust their behaviour in future attempts. This is particularly important if a mistake is made that may not affect a novice's perception of the successful completion of the goal of the procedure, but might cause long term harm to a patient. One of the main disadvantages of physical simulations is that it can be difficult to extract performance feedback from the model as Burdea *et al.* [16] note when comparing physical models of a human prostate to Virtual Reality models. Rubber prostate models were judged to provide a better feel for a real prostate than a PHANToM largely to the lack of tactile information provided by the PHANToM. As such, small local surface features were difficult to distinguish. The rubber models have however proved difficult to use as it is difficult to alter the model for different conditions, and there is no method for extracting performance feedback from a doctor other than through the final diagnosis.

The majority of the virtual medical training simulations developed do address this issue of feedback by analysing the procedure data and presenting it to the user after he or she is finished. For this to be achieved, there must be a goal for the procedure defined, as well as an optimal method that can be compared to the current user's procedure. In fact, one important advantage of a virtual training system is that a user can be given a standardised performance rating for the procedure performed based on the optimal method for the procedure. This could eventually lead to an objective method of certification of medical trainees or specialists [44]. Determining the performance in a medical procedure is difficult however, since it can be a complex, multi-dimensional task with many different outcomes – not just success or failure. Metrics will be dependent on the training task performed. Metrics may be defined for how close the simulated procedure came to achieving the goal of the procedure, but also for how that goal was achieved. Gorman *et al.* [42] suggests the following metrics for a task involving driving a simulated needle through a target overlaying a blood vessel: time on task, accuracy, peak force applied, tissue damage, surface damage, and angle error. However, they note the difficulty in calculating tissue or surface damage accurately. An argument can be made that time on task is not an appropriate metric when taken by itself, as it takes no account the success or failure of the procedure, or damage caused during the procedure. However, when considered with other metrics it may give a measure of experience and confidence. An experienced surgeon who has performed the procedure many times would be expected to complete the task successfully faster than a competent but inexperienced surgeon.

For a palpation simulator where the user may wish to examine the whole of an object for specific shape or surface properties, accuracy and angle error may not be so relevant. Particularly in training for diagnosis, metrics can be very high level. For example, in a large animal ovarian simulator, users palpate ovaries for a follicle to diagnose the stage of ovulation of the animal. The users might be asked "Does a follicle exist on either ovary, and if so, what size is the follicle". More difficult to measure - but equally important - are metrics such as search strategy used or thoroughness of exploration that may also be tested.

Systems exist to allow user performance to be stored over time [14]. This history feature allows any trends of improvement or otherwise to be noted. Persistent flaws detected by a system such as this can also be noted and addressed.

As well as training purposes, simulators should be considered from the point of view of assessment. The perceived skill level of a clinician for a procedure is often based on the number of times he or she has complete the procedure. This however may not always be the best method. Royston [97] suggests the eventual certification of surgeons through simulator testing. With the rich feedback that can be gained through these systems, the potential is there to assess individual aspects of an examination against an optimal examination. Certification can be seen as a method of proving competence, which is becoming increasingly important for hospitals from the point of view of defending themselves from litigation.

Michell [77] argues the need for both the medical and veterinary professions to introduce the practice of regular reassessment and revalidation. He notes that the General Medical Council now expects a "*regular demonstration by all registered doctors that they remain fit to practise in their chosen field*", and again draws parallels with the aviation industry where this has been standard practice for many years. Revalidation is particularly important in a profession where medical personnel are required to keep up to date with current practices, and operate within the limits of their training. Michell suggests that simulation might provide a solution in that it offers a potential method for providing analysis of performance and tutoring. He notes however that the set up and running costs for such a scheme would currently be prohibitively expensive.

The systems described in this section all provide post procedure analysis that will allow a user to adjust his or her technique the next time that the procedure is attempted. However, none take advantage of the possibility of using the advantages of a Virtual Reality environment to offer guidance or support to the user during the procedure to allow him or her to immediately correct his or her behaviour. Users of a simulator might never have tried the simulated procedure before, or might be experienced surgeons trying to refresh their knowledge of the procedure. A Virtual Reality simulator could offer the possibility of presenting a user with a level of feedback and guidance depending on his or skill at the procedure.

3.6 Conclusions

This chapter has presented an overview of some of the issues involved with Virtual Reality in medicine, and in particular surgical simulations. Previous examples of work presented in the literature have been discussed in order to highlight current areas of research, and suggest possible areas where more work is required.

This chapter relates to all three key research questions introduced in Chapter 1. The first thing that can be concluded from this chapter is the lack of related research from a veterinary point of view. There are many examples of human medical simulators, but currently no examples of veterinary training simulators in the literature. This is one factor that should be addressed. Very few of the previous examples have actually addressed a simulator as a teaching tool that is to be integrated into the training of a novice doctor. Most of the research presented has concentrated on the development of high fidelity simulations without addressing how they would be used.

There is some evidence to support the success of medical simulators as a training tool. In particular, one study with the MIST simulator has suggested that there are performance improvements from using this simulator on physical simulations that could be applied to real life procedures. There is literature to support that simulation can aid novice surgeons when learning the complex mappings required to operate MIS tools. The majority of the validation work has concentrated on this fact. There is less evidence to suggest haptics can be used as a tool for training in palpation procedures. In general, the studies presented in this paper do not provide solid evidence required to allow a simulation to become an accepted part of a training course. The ethical issues involved with validation of simulators have made this a difficult area to address.

Performance feedback has been discussed as one important advantage of simulators. This is one particular area where simulators can provide feedback that is not generally possible during traditional training. The use of performance feedback to train and assess medical personnel has been discussed, although there has so far been little work in this area.

This chapter and the previous chapter have given an overview of the current state of the research in the areas of haptics, and in Virtual Reality in medicine. The following chapters will build on this work to describe the design and evaluation of a veterinary palpation training tool.

4 Design and Development of a Palpation Simulator

4.1 Introduction

This chapter will examine the factors involved in the integration of a simulator into the Glasgow Veterinary Course. It will therefore look to answer the first of the key research questions stated in Chapter 1. The designed process heavily involved experienced veterinarians from the Glasgow University Veterinary School. The goal of this work is to build a tool to provide training to veterinary students in large animal ovary palpation techniques. Although the design of the system is biased towards horse ovary palpation, it is important to note that similarities exist between the problems facing universities in teaching equine and bovine ovary palpation.

4.2 Background

The training tool will take the form of a Virtual Reality environment that allows the students to practice palpation on virtual horse ovaries. However, it will only prove useful if can be integrated with the current horse reproduction syllabus of the veterinary school. Accepted computer aided learning (CAL) techniques will be employed to design the simulator such that it contributes to the course. The CAL design technique selected is the ABC method [79] developed at Glasgow University. This method has been shown to provide teaching material of a higher relevance than traditional methods. It is based on Laurillard's conversational framework for higher education learning [65]. The simulator will be designed to integrate into the Glasgow University Veterinary School course. However, courses vary throughout different universities depending on the resources available so the simulator may not be suited to every veterinary course.

Neufeld [85] encourages course designers to be wary of the side effects of introducing a new training tool or paradigm into a course. This new tool will almost certainly be introduced to the detriment of teaching other areas of the course either by replacing a lecture or tutorial that currently exist in the course, or by increasing the workload for the students. Most importantly, the tool must address some area of the course where, possibly through lack of resources or ethical considerations, there is a need for a different form of teaching. Use of the tool should complement the other relevant material presented in the course to give a more complete view of the problem than would otherwise be possible. It is important therefore to use a proven design method to ensure a training tool that will be more relevant when integrated into the course. Arthur *et al.* [2] present a convincing argument for the use of an accepted design methodology. Lack of contact with the user group during their design process led to a MIS simulator with no force feedback provided through tools, where as these haptic cues are often used by surgeons to navigate during procedures.

4.3 Activity Based Computer Aided Learning

The Activity Based CAL (ABC) method is a CAL design method developed at Glasgow University. It is based on the conversational framework developed by Laurillard [65] - shown in Figure 16 - that provides a teaching model for higher education. The conversational framework incorporates aspects of both the instructivist and constructivist educational models, and has a wide level of acceptance in higher education.

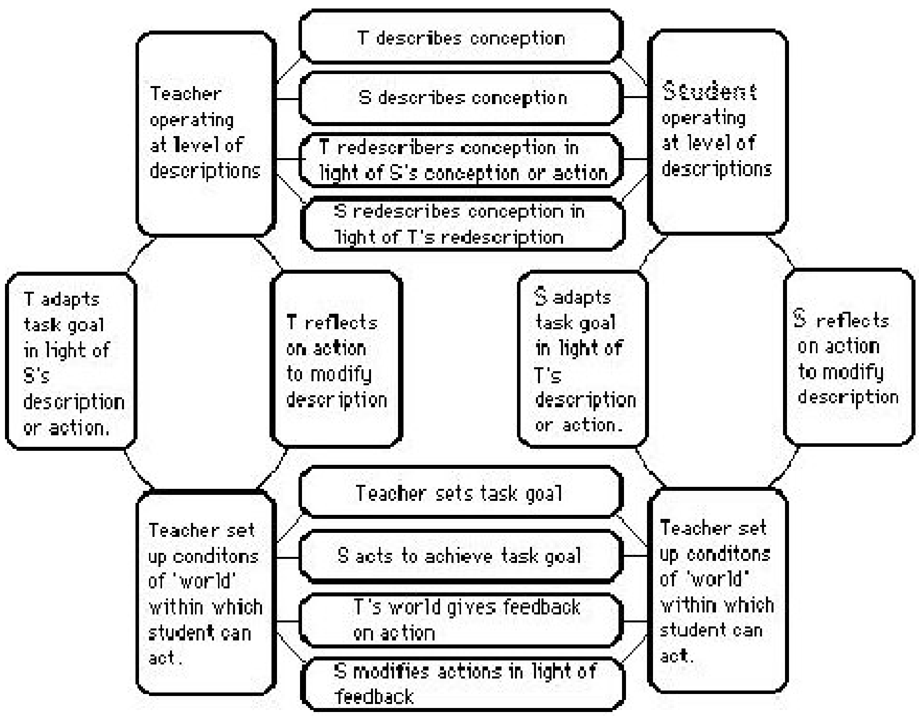


Figure 16. The conversational framework educational model for higher education proposed by Laurillard [65]. It consists of 12 stages of interaction, and is based around a dialogue between teacher and student.

There are twelve stages described in the model. These form the basis of a dialogue between the teacher and student to promote learning. These stages are:

1. The teacher (T) describes a concept to the student (S)
2. S describes the concept to T
3. T redescribes the concept in light of S's description
4. S redescribes the concept in light of T's description
5. T sets a task

6. S performs the task
7. T provides feedback on the task to S
8. S modifies actions in light of T's feedback
9. S reflects on the action to modify the description of the concept
10. S adapts action in light of T's description
11. T adapts the task goal in light of S's description
12. T reflects on action to modify the description

This approach acknowledges that any single teaching event may not reliably cause learning. It ensures that a concept is explained through both description and action. Also, the discussion is such that the teacher is aware of any misunderstanding to the student and can modify their description accordingly.

4.3.1 The Activity Based CAL Design Method

Activity based CAL (ABC) uses the Conversational Framework in the design process for CAL material. It has previously been used before to provide benefit in the design of a CAL system [80].

ABC can be carried out in six stages:

1. State the aims and objectives of the teaching. This can be done through a formal requirements capture.
2. Using the conversational framework, decide how each of the activities will be supported. This will be either through human-human interaction, human-computer interaction or by some other means.
3. State the resources that will be needed to support the activities in the mode stated in Stage Two.
4. Implement the design.
5. Evaluate the design. This can be done through user testing or by other means. Also, each activity can be compared against the conversational framework to detect any deficiencies.
6. Redesign, re-implement and re-evaluate the system as necessary.

This iterative approach ensures that each of the initial objectives is included in the teaching to as great an extent as possible. It is important that a piece of CAL software integrates into an existing course and by using a tested CAL technique such as the ABC method, we can ensure that this is the case.

4.4 Requirements Capture

Formal interviews have been conducted with three veterinarians who are involved in teaching the equine reproduction course at Glasgow Veterinary School. The purpose of these interviews is to gather information about the course, and methods of teaching. This information will lead to the development of requirements the virtual simulator must fulfil to be beneficial to the course. The interview consisted of the same thirteen questions asked of each participant. These questions are shown in Table 1. However, answers provided by the participants often prompted a discussion in which extra questions were asked. The questions were designed to gather information on the structure and content of the equine reproduction course at Glasgow University Veterinary School. This is important as the simulator must integrate into the course, and provide teaching in an area that is difficult to provide training using traditional methods.

What are the course objectives for teaching equine ovary examination working through from 1st to final year?
How is each of the course objectives met?
How well can each of the course objectives be met?
How is time divided into lectures/tutorials/lab work/practical experience for each year?
What are the core skills involved in the procedure?
What is the hardest thing for the students to learn?
Is training commonly provided from outside university sources?
What part does self-learning through (library work etc.) play in training?
What other areas would you like to see included in the equine ovary palpation course objectives?
What training could be improved upon?
What would be your requirements for a simulation to be incorporated into the current course?
How do you feel a simulator could fit into the ovary palpation training course?
Can you think of anything else to discuss?

Table 3. The thirteen standard questions asked during each interview.

For two of the interviews, audio recordings were taken during the interview. The audio was reviewed at a later date, and statements gathered from the questions were recorded. Notes were written on any answers given during the third interview. All three people interviewed were familiar with the PHANToM force feedback device [72].

To analyse the interviews, five categories were created to sort the relevant quotes from the interviews. These are shown in Table 4. These categories were chosen such that information would fall into only one category in as many instances as possible. These categories were chosen to provide information for use with the ABC method. ‘Reproduction Course Information’ and ‘Perceived Course Deficiencies’ provide an insight into the current course, and ascertain the perceived weaknesses. This information will be used to build a list of aims and objectives for the course that are necessary for the ABC CAL method. The ‘Simulator Information' section provides insight into the experts’ perceived potential benefits of a training simulator. This will be used when designing the practical aspects and functionality of the simulator. The ‘Ovary Examination Procedure’ information will provide information for the models used in the simulation. This, along with other sources, will also be used to develop an understanding of the procedure being modelled. ‘Other Information’ is a general purpose category designed to capture all other useful information provided by the experts that does not fall into any of the other categories.

Reproduction Course Information	Any information on the **current content** or **objectives** of the **equine reproduction course**, or **known changes to the course**. Also, statements about **the teaching methods** of the course, and **methods of learning** employed by the students. Statements providing information about **student numbers** on the course.
Perceived Course Deficiencies	Any deficiencies perceived in the course material or resources, and suggested changes for improvement in teaching or training that are not currently being implemented.
Simulator Information	Statements about the **potential use** of virtual simulator training. Any **requirements** for a simulator, and statements on how it would **integrate** with current teaching methods.
Ovary Examination Procedure	Any information specifically relating to the **practical aspects** of performing ovary **examination** or **manipulation**.
Other Information	Statements providing information that is **not covered** by the above categories.

Table 4. Categories for analysing information collected during the interviews. The words in bold were used to emphasise the keywords for each category to the independent reviewer.

A sample of the transcript, which was a subset of the full transcripts, was given to an impartial reviewer with no connection to the project. This reviewer was asked to divide the transcripts into separate facts and place each of them into one of the categories above. It was emphasised that the reviewer should split the dialogue into small single facts such that no single fact returned by the reviewer could be divide into two different facts that may possibly be in two different categories in the initial analysis.

The independent reviewer returned 279 facts in total, separated into the five categories. These facts from the sample transcript were compared to the full analysis to check for agreement. There were 25 facts that were placed in a different category by the independent reviewer indicating that over 90% of the reviewers facts were placed in the category expected by the initial analysis. In no cases did a fact from the reviewer overlap with facts from two different categories in the initial review. Facts that were placed in differing categories by the reviewer from the initial review are shown in Appendix A.3.

Each of these statements made by the different interviewees was compared to other similar answers to search for both agreements and differing points of view. The results of this analysis were used to build up an accurate picture of the veterinary reproduction course, and requirements for a simulator that would integrate into the course. These results are now discussed in further detail in the following sections. Unique codes have been given to each of the quotes, and these can be viewed in Appendix A.3. These codes are used to provide reference to a quote that is used to back up the relevant statement. The initial part of the code indicates the category that it has been sorted into. These are:

- CI - Reproductive Course Information
- PCD - Perceived Course Deficiencies
- SI - Simulator Information
- OP - Ovary Examination Procedure
- O - Other Information

The digits following indicate a unique number for the quote inside the given category

4.5 Horse Reproduction Course

Veterinary students must learn both the background knowledge, and the physical skills required to perform the various procedures. This is reflected in the aims of the course as well as how the course is taught. Lectures supply most of the background knowledge. The physical skills are taught mainly through in-vitro palpation (in anatomy labs), and to a lesser extent *in-vivo* palpation with animals at the veterinary school. The veterinary course at Glasgow University is a five-year course. It is designed to cover a range of techniques relating to both large and small animals.

4.5.1 Course Division

The course division information has been built using comments from all three interviews. During the course of the interviews (held early in the year 2000), it became apparent that the course was undergoing changes during the 2001-2002 year. The course at the time of the interviews is described first, then the changes to the course are described. The source of the course information is indexed with respect to the numbered quotes in Appendix A.3.

During their first year, students will not directly be exposed to any material from the horse reproduction course. However, gross equine anatomy taught in first year is relevant to the reproduction course taught later (CI 17, CI 63, CI 70). This is taught through lectures and is taught at a superficial level at this stage in the course (CI 18).

In second year, students will start to learn about the physiology of the equine reproductive system as well as the anatomy (CI 19, CI 62, CI 70). The physiology material concentrates on the hormonal events that control the cycle. This year's reproduction material is taught through two lectures on male and female equine reproduction and participating in one anatomy lab (CI 3, CI 9, CI 62). In this anatomy lab, students will be presented with a bovine tract placed in a tray for examination (CI 63, CI 66, CI 67). Some of the lecture material taught is generic, and students would be expected to extrapolate between animals (CI 18, CI 80). At the end of second year, students should know the structures involved in equine reproduction, and their functions (CI 20).

During the students' third year, there is very little course information on equine reproduction. They are presented with minimal material on medication that contains some related aspects (CI 21, CI 22). This is, however, the only third year subject that has any relevance (CI 22). This material is all taught through lectures (CI 40, CI 78).

In the fourth year, the students are presented with more in-depth material, the main focus of which is fertility management. This would include the manipulation of the reproductive cycle through the use of drugs. Also, early pregnancy diagnosis, infertility investigation, and ultrasound scanning in examinations would be taught (CI 24, CI 25, CI 26, CI 71). This is the main focus of equine reproduction work in practice. Fourth year consists of eight lectures on male and female reproduction, although this figure will soon be reduced to four lectures (CI 23). This is the stage of the course where the material taught in first and second year is presented in more detail (CI 23). The fourth year equine reproduction material is all lecture based, although these lectures are reinforced with slides and videos (CI 41, 43). The lectures come from an outside source, a veterinarian who specialises in teaching this material (CI 49).

The final year of the course provides a much greater emphasis on practical material, designed to provide an environment that is closer to veterinary practice. Practical training sessions with demonstrators are reinforced with simultaneous ultrasound and palpation examinations (CI 44). In the current scenario, the veterinary school try to allow every student to have minimal experience of *in-vivo* horse ovary examinations (CI 3, CI 12, CI 69). This is an achievable goal, but only because expectations for the students proficiency levels are not particularly high (CI 47, CI 32, CI 37, CI 38).

Cost implications and ethical considerations do limit the opportunities for students to palpate ovaries (CI 33, CI 34).

In the course described above, every student who progresses through the veterinary school is exposed to the same modules. Since these interviews have taken place, changes (referred to as ‘student tracking’) have been introduced into the course (CI 59, CI 75). The experts interviewed were aware of the changes at the time of the interviews, so were also able to describe the new course divisions. The course remains unchanged for first to fourth year students. However, final year students will now be able to choose subjects for specialisation. These subjects are small animal, equine, and farm animal (CI 84). The final year students now participate in core modules for both the equine and farm animal courses, but must also choose to specialise in one or the other (CI 83). This has been introduced specifically to reduce the number of students participating in the modules, such that there will be a higher ratio of animals to students, and therefore more opportunity for the students to gain practical skills. This will allow more opportunity for those interested in equine work to gain experience in horse examinations.

4.5.2 Additional Teaching Resources

Multimedia CDs and videos of recorded postgraduate lectures, some of which are relevant to the equine reproduction course, are also available for students who are keen to learn more (CI 45, CI 52). These teaching aids are direct digitisations of postgraduate lectures. They provide an audio recording of the lecturer's voice, while showing the corresponding slide from the lecture. The CDs take no advantage of other potential benefits of multimedia, such as the use of hyperlinks to present or navigate through information.

The veterinary school has recently purchased a fibreglass model of a horse (O 19) that is shown in Figure 17. This can be used in conjunction with a preserved specimen tract to provide a simulation of internal palpation. In particular, this device is used to teach students intestinal palpation using preserved horse intestines in different configurations. The main function of this training tool is to teach students to diagnose colic, which is a distension of the intestines caused by excess air, water or solids. Despite its advantages, it is still not in general teaching use (CI 10). This is possibly due to the large set up time, the fact that only one device means that only one student can use it at a time, and the requirement for a tutor to be present during the teaching. The best method outside of *in-vivo* training tends to be feeling tracts in the anatomy lab while looking away (CI 11). However, these tracts will be bovine and not equine tracts. The ovaries for the fibreglass horse are plastic. They have an appropriate shape and positioning within the horse, but there has been no attempt to provide a realistic feel for the ovaries.

Figure 17. The fibreglass horse used to train students in palpation of horse intestines.

A library is also available at Glasgow University Veterinary School for use by the students (CI 77). Students would be expected to use the library to find out additional information when necessary, but there is only so much that can be learned from reading books (CI 7). A computer cluster is also available for student use (CI 77). The Computer-aided Learning In Veterinary Education (CLIVE) project is being developed at Glasgow University Veterinary School as an internet based resource for veterinary student education. This takes advantage of multimedia technology to allow students to learn through text, pictures, animation, and interaction with the system. Students might also learn through talking with experienced vets (CI 9).

Students who are particularly interested in equine work may get experience from work placements in equine veterinary practices (CI 51). Approximately five percent of students at Glasgow University Veterinary School take advantage of this to gain experience in equine reproduction (CI 50). There are, however, still problems that restrict a student when gaining practical experience through work placements (CD 17). This is discussed more in Section 4.6.1.

4.6 The Horse Ovary Palpation Procedure

The dorsal view of the female equine reproductive system is shown in Figure 18. During an ovary examination, the veterinarian inserts a gloved hand into the pelvic area of the horse through the rectum. The veterinarian must search through the pelvic region of the horse for the uterus. The ovaries are attached to the uterus, and each can be found by following either the left or right uterine horn. This is difficult in itself, since the veterinarian must perform this search through touch alone. It usually requires several attempts before an inexperienced student can locate an ovary. Once located, the veterinarian will cup the ovary with one or more fingers, and palpate it using their thumb. In particular they will look for any abnormalities in the shape or surface properties of the ovary, and through experience, will be able diagnose different conditions through touch alone.

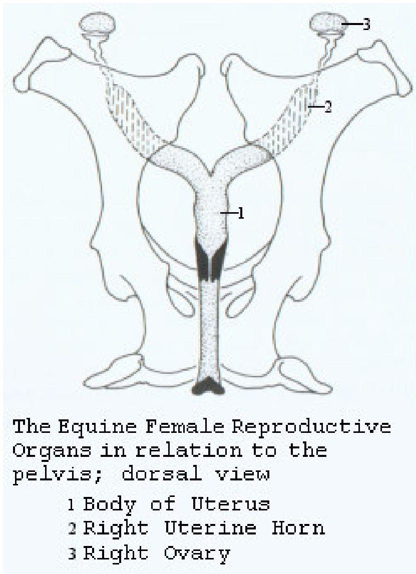


Figure 18. Dorsal view (from above) of the reproductive system of a non-pregnant mare from Dyce *et al.* [32].

A common task carried out by a veterinarian is to locate follicles on the surface of the ovaries. A follicle is a spherical fluid filled sac that grows on the surface of an ovary some of which exists under the surface. A follicle will typically grow from a small size to a few centimetres in diameter. As the follicle grows, it will also tend to move towards the centre of the ovary. Depending on the size, position and feel of the follicle a veterinarian can diagnose the stage of ovulation of the horse. There may be many follicles on each ovary, but only one follicle will eventually ovulate. Other features may exist on the ovary surface such as a corpus luteum which may be mistaken for a follicle by novices. Unlike a follicle however, it is ridged around the edge. Ovarian cysts or tumour may also exist, but are less common. Typical descriptions gathered from one interviewee are:

An ovary is a hard fibrous object. It may feel similar to some objects in the abdomen and is therefore difficult to identify (OP 39).

A corpus luteum is a surface feature on an ovary that may feel similar to a follicle. However, unlike a follicle, it has a palpable thick walled ridge (OP 37).

A follicle is a thin walled, soft partially submerged object on an ovary's surface (OP 38).

The core skills involved are location and identification of the ovaries, and recognition of surface features on the ovaries. Ultrasound scanning is now often used to identify the surface features.

However, both ultrasound and palpation techniques are considered relevant and are still taught by the veterinary school.

4.6.1 Course Deficiencies

It is clear from the transcripts of all 3 interviewees that there are deficiencies in the current Glasgow veterinary course. Both time constraints and limited financial resources restrict what can be taught to the students. From the point of view of equine ovary palpation, the largest deficiency is in providing opportunities for students to gain practical experience (CD 1, CD 9, CD 14). Glasgow Veterinary School owns a small herd of ponies. These ponies can be used to provide practical experience to students, but a herd of ponies is expensive to maintain (CD 5). Animal welfare considerations restrict the number of invasive procedures carried out on a horse, and indeed guidelines do exist that restrict the number of times an animal can be rectalled in one day. In the interest of fairness, similar training opportunities must be given to all students on the course. In some ways, providing a small herd of ponies can be almost worse than having none, because students who do not get the opportunity to gain experience in equine rectalling due to the low ratio of ponies to students will feel cheated (CD 6). The Veterinary School will also at any time have a number of horses that are being treated by the school clinicians. Again, there are similar problems in using these horses for invasive training, as they belong to owners outwith the veterinary school who are unlikely to agree to allow students with little or no experience to practice on their horses.

Students are expected to participate in training outwith the veterinary course. Approximately five percent of all students will get placements in a specialist equine practice. This is one of the most useful ways of gaining experience in equine medicine. However, particularly for a procedure like ovary palpation, a student may still not get a chance to examine a horse (CD 17). Again, a horse owner is unlikely to allow an inexperienced student to examine his or her horse as there is the possibility of injuring the horse. A newly qualified veterinarian may therefore be required to learn ovary palpation in the first weeks of his or her job (CI 6). This is far from an ideal situation and could lead to costly mistakes being made.

Currently, students spend very little time in the anatomy laboratory learning about equine reproduction. There is also a long period of time between the practical lab in second year, and the in-vivo experience that students may get in their final year. It is doubtful that the students will be able to maintain the skills obtained in the second year lab through to the final year session.

4.6.2 Suggested Improvements to the Course

The interviewees were also asked about improvements to the course that could be made if the extra resources became available. This was on top of the previously mentioned changes that have recently been introduced by allowing students to specialise in subjects in their final year, and therefore reducing the numbers on the different sections of the course.

As expected, each veterinarian interviewed had a slightly different perspective on how the course could be improved given extra resources, but one factor that all agreed on was the need to increase the amount of practical experience gained by the students during the course (CD 1, CD 12, CD 14). This is to be expected as examining a tract laid out on a table in an anatomy lab provides a different experience than examining the tract within an animal. When laid out on a table, the tract takes a different shape and the structures are in different relative positions than in an animal. Experience with animals, on top of experience with specimen tracts, is considered important to training students. It is this *in-vivo* experience that would be beneficial to the course if the resources were there to provide it (CD 15). It is not financially possible to achieve this by dramatically increasing the number of horses available to the students, so other solutions - such as teaching aids - must be looked at (CD 12). For the students who are interested in pursuing a career in equine veterinary work, it is important that the university provides them with methods to improve their practical skills to a higher level (CD 9).

In order to address the length of time spent between practical training, first in the anatomy lab and then in in-vivo training, one participant suggested more anatomy labs that last a shorter length of time. This would provide a suitable alternative that would allow repeated training sessions for students at more regular intervals (CD 3).

4.6.3 Aims

Currently, students graduating from Glasgow University Veterinary School are not expected to be able to fully perform a horse ovary examination. The expectation is that a student will be able to locate and identify the ovaries, without damaging the horse. This is done by recognising the size and consistency of the object. However, they will still require training before they can be trusted to perform examinations in practice, where a mistake can easily lead to the loss of a developing foetus. The aims from the course gathered through interviewing have been split into each of the five years:

By the end of first year, students must be able to describe of the anatomy of the equine reproductive system. This is used as a basis for future ovary palpation (CI 17, CI 19, CI 62, CI 78).

By the end of second year, students must be able to describe the physiology of the reproductive system. This would be at the level of knowing the fundamentals of the anatomy and the hormonal events that control the fertility cycle. This is taught largely generically, and students would be expected to extrapolate between species where possible (CI 17, CI 19, CI 80). By the end of second year, students should be able to describe structures and their functions. But not in any great detail (CI 20).

Third year contains little information relevant to the equine reproduction course. Students however must be able to describe methods of manipulating the reproductive cycle (CI 21, CI 22, CI 23).

In fourth year, normal cycles and manipulation of cycles through medication are addressed in depth. Students must be able to describe fertility practices such as artificial insemination (CI 22, CI 23, CI 24).

There are no specific aims for training students in the equine reproduction material in fifth year. The Veterinary School tries to allow all students to perform one or more in-vivo equine ovary palpation, but no skill level needs to be reached in order to pass the course.

4.7 The Use of a Virtual Reality Simulation for Teaching

An important feature of a training simulator for students (if it is to be accepted is into a course) is that it must be able to integrate with the course. If the resources are not present to make use of the simulator then it will not be used. The Glasgow University Veterinary School has adjusted it's course (as described above) to reflect the fact that not all students will require equine veterinary skills in their working life. There is no need therefore to develop a simulator to accommodate all students in the course, only those interested in equine veterinary medicine (SI 4). One veterinarian interviewed suggested problem solving tutorials would be one teaching situation that would benefit from simulator training (SI 18). These could take the form of scenarios that a student could work through from start to finish, incorporating all aspects of the examination. This would be particularly useful for final year students (SI 17). The fibreglass horse model shown in Figure 17 is not used to its full potential because of the time it takes to set up, as well as the fact that a demonstrator can only teach one student at a time. This would seem to be a flaw with any simulator developed. The number of students using a virtual simulator at the one time is limited to the number of haptic devices available. However, a simulator should offer the possibility of providing a method of self-assessment to students (SI 12). If a simulator was included in a software package specifically designed for training, and self-assessment, a demonstrator would not need to be present.

All of the veterinarians interviewed felt that a simulator could be fitted into the course. For example, for students particularly interested in equine work, a simulator session could be included at the expense of a sheep lab session (SI 3). A simulator was not seen as a replacement for traditional training methods but as supplemental to the current lab and tutorial sessions (SI 6, SI 15).

One of the key aims identified by all the interviewees for the simulator is flexibility (SI 1, SI 7, SI 8). Currently, anatomy lab techniques are used to provide the majority of ovary palpation training. However, they tend to lack realism in terms of relative position of structures on the tract, and students are only exposed to the cases that the specimen ovaries allow. When a condition is being taught in the lectures, students may not be able to get experience in recognising or treating the condition if the veterinary school does not have an animal with the condition. Surface features on ovaries will change in position, size, and frequency, and demonstrating this using specimen ovaries is not always possible. Specimens decay over time, and students will only be exposed to specimens that are available during their anatomy lab sessions. Haptic properties might also change as the features become different sizes. For example follicles start small and hard, but grow softer as they get bigger. Tumours can vary widely in feel. A simulator should allow a lot of flexibility in the models developed. This would extend to allowing students to experience immediate changes in the structures that may take several days to change in real life (SI 11). An example of this would be for students to experience palpating mares moving through from early follicles to pregnancy diagnosis. This could all be achieved in one sitting at the simulator where as several trips to a stud farm over several weeks would be required in the real life situation.

One interviewee was keen on scenario based learning (SI 7). That is presenting the student with a medical history of the animal and a condition, and allowing them to examine it and make a diagnosis. It is more likely that a veterinarian will deal with this situation in actual veterinary practice. This would be particularly useful for final year students who will soon be expected to be able to carry out consultations with clients (SI 17).

All interviewees stated that realism was important (SI 2, SI 9, SI 10). Students must be able to relate experiences on the simulator to real life situations. To some extent, the simulator offers the potential of greater realism than the anatomy labs, as it will allow physiological changes to be modelled as well as anatomy. As the veterinarians interviewed had all tried an initial version of a virtual ovary palpation simulator, they were aware of its current design. They suggested the possibility of including extra environment features would greatly increase realism (SI 14). It is important to note however that absolute realism was not viewed as a requirement of the simulator (SI 5).

A simulator should also take advantage of features that are not available in traditional teaching environments. Equine ovary palpations are performed without the aid of visual cues. This is not only true for the students, but also for the teacher who relies on the student to describe what he or she is feeling (SI 20). A simulation should allow a teacher to observe a student's examination. This could provide a powerful teaching tool (SI 19). Other features could be integrated into the simulation such as ultrasound pictures. This would provide a student with visual and haptic representations of the ovaries as well as the corresponding ultrasound picture (SI 16).

4.8 Teaching Ovary Examination

When examining the aims for the ovary examination procedure, there are three factors that are important to a veterinarian and are taught during the reproduction course:

- Veterinarians must understand the anatomy and physiology of the horse reproductive system;
- Veterinarians must be able to locate and palpate the relevant structures;

- Veterinarians must be able to use ultrasound scanning equipment to examine the relevant structures.

ABC activity charts for teaching equine ovary examination are shown in Appendix A.5. Looking at the ovary examination section of the equine reproduction course using the Laurillard conversational framework as a teaching model, the main weaknesses identified are in steps seven to twelve. Because of the small number of practical training sessions, and the time between sessions, students get few chances to repeat the task. They do not get a chance to modify their behaviour based on feedback received from a teacher.

Activity Chart B in Appendix A.5 demonstrates how a simulator would contribute to the course again using the Laurillard conversational framework. The key aspects for the simulator identified are that it must be flexible, and accessible. It must offer students different advantages to those of the anatomy lab. It must complement the training currently offered. Using the flexibility of a simulator, a number of cases could be modelled that students would be unlikely to see in the anatomy lab.

From Activity Chart B, we can see that step 5 in the conversational framework could benefit from simulator training. An advantage of the computer model over an anatomy lab or *in-vivo* training session is that a student can feel and see the anatomy and physiological changes of the reproductive system. The activity chart also demonstrates how steps 6 and 7 benefit from a simulator. In an *in-vivo* practical session, a teacher will not see or feel what the student palpates. The student must communicate to the teacher what he or she feels, and the teacher must interpret this and make an informed guess about the location of the student in the abdomen. A simulator can allow a teacher to see the student's examination and even interact with the student's cursor to guide his or her movement.

In steps 8, 9, and 10, a simulator can offer other benefits. It is clear from the interviews that *in-vivo* experience is an important part of training and a simulator will not be used as a replacement to *in-vivo* training, but as supplementary training. The main problem of the traditional lab and practical training sessions approach is that it is difficult and expensive to provide regular sessions for student training. A simulator must therefore be accessible to the students. If a simulator is to achieve the goal of accessibility, it must be available to students regularly. This would suggest that a system that could offer useful feedback to the user without a teacher present would be beneficial. The fact that the non-virtual simulator is not in general use emphasises the importance of this. Students cannot use the non-virtual simulator without a significant set-up time and qualified veterinarian or technician to aid them. A virtual training tool could be accessible all year round, and guidance and performance feedback supplied by the virtual environment would reduce the need for a teacher to be present.

Alternatively, the combination of a simulator and human tutor can provide another method of training and assessment. In a large animal ovary examination, the teacher cannot see the movements of the student. The tutor can watch the student practice on a simulator and provide feedback to the student either during or after the virtual examination.

Steps 11 and 12 of the activity chart show again that the flexibility of a virtual system could provide benefit. The complexity of the models and environmental features can be altered unlike in non-virtual tracts. In a virtual system that incorporates assessment of the student's performance on the simulator, automatic adjustments of the surrounding environment could be made to allow the student to explore a more challenging environment if he or she is performing well, or a less challenging environment if he or she is performing poorly. This is difficult to achieve in non-virtual teaching as teaching is restricted to the tracts and horses available at the time of teaching.

4.9 Conclusions

The interviews conducted with the vets involved in teaching the reproduction course at Glasgow Veterinary School confirmed that there are currently problems in teaching both bovine and equine ovary palpation to students. Lack of resources, and ethical guidelines along with other factors described above can restrict opportunities for students such that even newly qualified vets are still very inexperienced in this area. A Virtual Reality simulator teaching tool offers one potential solution to this problem. This section has investigated the requirements for such a training tool, and how a simulator training session could be incorporated into the course. Through the use of the ABC method, requirements have been gathered that will allow the training tool to complement the current course, and provide training in the areas where more is required. This study has examined the feasibility of simulator training as a component of the reproduction course. Alternatives to simulator training are expensive, and mean that significant annual costs must be met to maintain these resources. The interviewed veterinarians also confirmed that it would be possible to fit a session of simulator training into the course. This would be on the assumption however that there were enough devices to provide training for everyone in the tutorial group. These devices are currently expensive meaning high initial costs to set up a simulator tutorial, which may restrict the possibly of simulator tutorials. A simulator may provide more benefit if it is accessible to the students outwith tutorials. It should provide a learning environment in which it is not necessary for an experienced clinician to be present in order to provide the training. One method that this could be achieved is by providing the student with performance feedback based on his or her actions when using the simulator. This is difficult in ovary palpation even with a tutor present, as it is an internal examination, and feedback relies on the student describing to the tutor where they are in the environment. It should be flexible to allow a wide range of cases, particularly as reproductive conditions may be seasonal.

4.10 *Development of the Ovaries*

4.10.1 Background

There are previous examples in the literature of construction of anatomically accurate three-dimensional medical visual models [33]. A common method for building these visual models is to use a Computer Tomography (CT). That is, firing x-rays at a target from different angles in a plane to build a cross-sectional picture of the target. The CT scanner is used to take several cross-sectional pictures of the object to be scanned. These pictures can be integrated to form a three-dimensional virtual model based on an actual physical model. However, there is no analogous method currently available for scanning in the feel of an object.

There have been attempts to simulate specific types of tissue through complex measurement techniques, using known forces to deform the object. This technique is not possible for measuring properties of internal organs for ethical reasons. As the haptic properties of an organ may change between *in vivo* and *in vitro* palpation, laboratory specimen models may not provide the most accurate measure.

This section describes an attempt to approximate the feel of an equine ovary and three sizes of follicles. The technique described does not try to model the mechanical properties of the tissue, which could lead to complex and computationally expensive algorithms such as methods described in Section 3.3. An alternative method was used. Simple properties of the virtual objects were set by experienced horse veterinarians in order that the best approximation of the expert group participants' opinion - using only these parameters – was reached. A measure was also taken of their confidence that the virtual model was a good approximation of the physical object modelled. During this stage of the development, it is important to involve experts to provide face validity such that the virtual models look and feel similar to the objects that they represent. Involving experts in the design process will also lend credibility to the simulator, which is important if it is to become accepted as a training tool.

4.10.2 The Models

It was decided to concentrate on one stage rather than try to model the entire procedure. With the current state of haptic devices, it would not be possible to accurately recreate all the sensations involved in this procedure. Features such as the anal sphincter and the straining felt by the veterinarian on his or her hand would be difficult to model. The Fibreglass horse model described above can successfully provide students with the spatial layout of the internal reproductive organs of a horse. It was therefore decided to concentrate on the palpation stage of the procedure. As this is a difficult procedure to perform and opportunities to practice are limited, it is possible that many students will graduate without having palpated horse ovaries.

The models developed are simple geometric approximations of equine ovaries. An ovary is made by joining two distended spheres to make a bean shape object shown in Figure 1. These initial models were built in the early stages of the project with help from experienced veterinarians to adjust the shapes and sizes.

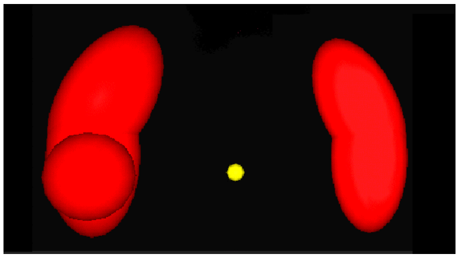


Figure 19 - The Horse Ovary Palpation Simulator. This environment consists of a left and right ovary. On the bottom half of the left ovary, a spherical follicle can be seen. The user's cursor is shown as the yellow sphere in the centre.

Although these are reasonable approximations of the physical object, methods have been explored to further increase the realism of their shape. An experienced veterinarian was given modelling clay to build a model of a horse ovary with no surface features. This model was scanned in using a three dimensional scanner at the Faraday Graphics Lab at Glasgow University to create a Virtual Reality Modelling Language (VRML) virtual horse ovary model and is shown in Figure 20. VRML is a standard language for modelling three dimensional shapes through polygonal models.

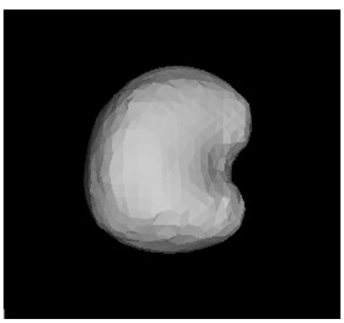


Figure 20 - A virtual horse ovary that has been created by scanning in a clay model at the Faraday Lab in Glasgow University. The clay model was created by an experienced horse veterinarian.

This was judged to be a much closer likeness to an actual horse ovary. However, when 6 expert veterinarians interacted with the model using the PHANToM all expressed a preference for the geometric models. This could be due to the fact that despite the standard GHOST polygon smoothing algorithm being used, the individual facets that made the VRML model could still be detected. Therefore, for the purposes of the experiments contained in this thesis, the geometric ovary model was used.

Deformable visual models were not seen as a requirement for the simulator. The users of the simulation would not receive graphical feedback from the models. For simulations involving one point of contact with the environment - such as interactions using a PHANToM - the object can often be modelled through objects that feel soft but do not actually deform. This can be done my increasing the force felt by the user more slowly as the user moves further inside the object. No other interaction point exists in the environment to detect any change in shape of the object. This is only possible as the virtual ovaries are fixed in space, and do not interact with the other objects in the environment.

4.10.3 Haptic Properties

4.10.3.1 Haptic Device

From Chapter 2, it has been argued that current force feedback devices can more realistically simulate three dimensional virtual objects than purely tactile devices. For this procedure, a veterinarian is interested in locating objects on the surface of an ovary. These objects will change the shape of the ovary, and may have different haptic properties (for example they may feel softer). The device should also allow the user to attempt to size the virtual ovary. Therefore, the important EPs for this technique are pressure, lateral motion and contour following. If location and identification of ovaries were issue for the simulator, the enclosure EP may also become important to allow the user to quickly get an overview of the shape of the object grasped, and identify it. However, the simulator designed will concentrate on the palpation stage of the procedure. From Chapter 2, it can be seen that the PHANToM is one force feedback device that supports the appropriate exploratory procedures. Unlike the HapticMASTER, the PHANToM also has a standard attachable thimble that should allow a more natural form of interaction than grasping a pen for a palpation procedure. One factor that must still be taken into account is that global shape and size properties are optimally judged through enclosure. As the PHANToM does not allow the enclosure EP, perception of these properties will be limited although still possible.

One disadvantage with the PHANToM for this procedure is that it is a single point of contact device. Even though a veterinarian may only use one finger to explore the surface of the ovary, he or she will not be able to grasp the ovary while doing this. A force feedback glove such as the Rutgers Hand Master would allow a user to grasp an ovary. However, the user would be restricted when exploring the ovary surface to whole hand exploration as this device only supplies one degree of force feedback to each finger.

There is currently no ideal device for haptically interacting with a virtual environment. There will always be compromises made as the current devices do not support all forms of interaction. It was felt that the PHANToM would offer the best compromise for this situation as it offers high resolution force feedback to one finger. It was important however to remain aware of the potential problems of restricting a user to one point of contact with the environment.

4.10.3.2 Methodology

As Arthur *et al.* state [2], consultation with an expert group is essential during the design process to avoid errors in the design that will make a simulator unsuitable for training in the task. This section describes an experiment that involves experienced large animal veterinarians in designing the haptic properties of the Virtual Reality ovary models used in HOPS. The experiment described in this section is designed to gather values for haptic properties that will provide an approximation of the feel of each of the ovaries and follicles being modelled.

The five participants involved in the experiment were veterinarians from the Weipers Centre for Equine Welfare. They were split into two groups (of two and three subjects). Each group performed the same task.

Although all were regular computer users, none of the participants had any previous experience using the PHANToM force feedback device. The participants were therefore allowed to familiarise themselves with the device. The standard Multi-3D demo (shown in Figure 21) included in the GHOST toolkit was presented for the subjects to learn to interact with a three dimensional environment using the PHANToM. They were then presented with the experimental environment, and encouraged to adjust each parameter to try to familiarise themselves with its effect on a virtual object.

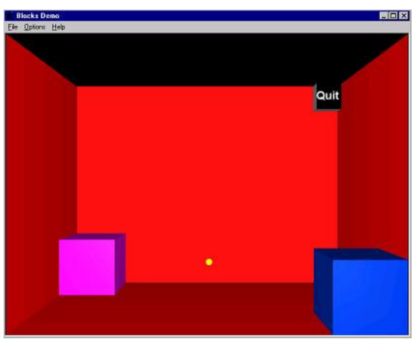


Figure 21. The standard Multi-3D demo distributed with the GHOST toolkit.

The experimental environment was the Horse Ovary Palpation Simulator developed at Glasgow University. It consists of two virtual horse ovaries fixed in a box in space. Spherical follicles could be placed on either of the ovaries.

The participants could adjust the feel of an object by adjusting three independent parameters - stiffness, friction and damping. These are the standard haptic properties parameters used by the GHOST toolkit developed by SensAble Technologies. Each of these parameters could be changed by adjusting the position of a slider. A textual description of the parameters was also provided for each subject. These descriptions are given below.

Stiffness – This controls how rigid an object is. A low value will give a soft object, like a grape, where as high values will give a rigid object like a block of metal.

Friction - This is the surface friction applied when moving over an object's surface. A high value will provide greater resistance to movement, where as a low value will provide little resistance. For example, sandpaper would have very high friction where as ice would have very low friction.

Damping – This controls how viscous the surface feels. A high value will feel viscous and sticky, where as a low value will not. A high damping value might feel like moving through syrup, where as a low damping value might be like moving through water.

The stiffness model used was the linear Hookes-Law model that is standard in the GHOST toolkit. This takes the form of

Reaction Force = constant (K) * penetration distance

The K value was adjustable between 0.01 and 1.0 with increments of 0.01.

Friction grouped together the static and dynamic friction components of the GHOST toolkit. This kept the resistance continuous such that there was no 'stick slip' sensation allowed. Friction was variable between 0 (no friction), and 1.0 (high friction) with increments of 0.01.

The damping model used was the standard GHOST model. Values ranged between 0 and 0.003 with increments of 0.00003. Values higher than 0.003 were found to result in increasingly stability problems with the PHANToM and therefore were not considered suitable for this experiment.

Five radio buttons laid out horizontally were provided for setting the level of confidence that the virtual model was a good approximation to the physical object. The description provided to each participant at the start of the experiment is given below.

Confidence Rating – This sets how confident the participant feels that the object being modelled is accurately represented by the model.

The furthest left radio button was marked as low, and the furthest right radio button was marked as high.

The task set for the participants was to adjust the three parameters to obtain the closest approximation to the feel of the requested object. They were encouraged to discuss the haptic properties within their group, and come to a consensus on the final haptic properties of the object. The objects used in this experiment were a left and right ovary, and three follicles of diameter 2cm, 3cm and 3.5cm. Each group was asked to set parameters for each of the ovaries once, and each of the follicle sizes twice. For each group, the ovary properties were set before the follicle properties as the follicles are partially submerged in the ovary. The follicles were presented to each group in a counterbalanced order with all haptic properties being set for the three follicles before the repeated follicle cases occurred.

4.10.4 Results & Discussion

Results from the analysis of the stiffness data are shown in Figure 22. It can be seen that the stiffness values returned for all of the follicles were similar in value. Considering the range of k values possible was between 0 and 1, both groups in all follicle cases used values in the lowest quarter of the range. An average of these values would return a similar stiffness value in feel. There are large differences however between the values set by Group 1 and Group 2 in the stiffness of an ovary. Taking an average here would not necessarily return an appropriate value. In order to resolve this discrepancy, both ovary models were presented to a separate horse veterinarian to compare. He stated that the stiffer model was a better representation of an ovary. This will also provide a detectable difference in stiffness between the ovary and follicle.

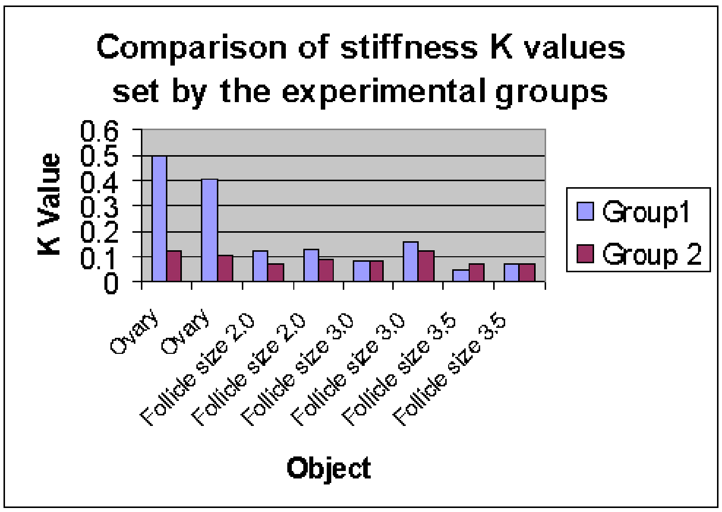


Figure 22 – Chart showing the comparison of stiffness k values for different objects between the experimental groups. Each object was presented twice.

The frictional values returned by both experimental groups are shown in Figure 23. All values were again far to the left of the available range. Again the largest difference between the groups was in selecting the value for an ovary. Again both cases were presented to a separate veterinarian. He reported that there was no discernible difference between the two values, and that both would be appropriate.

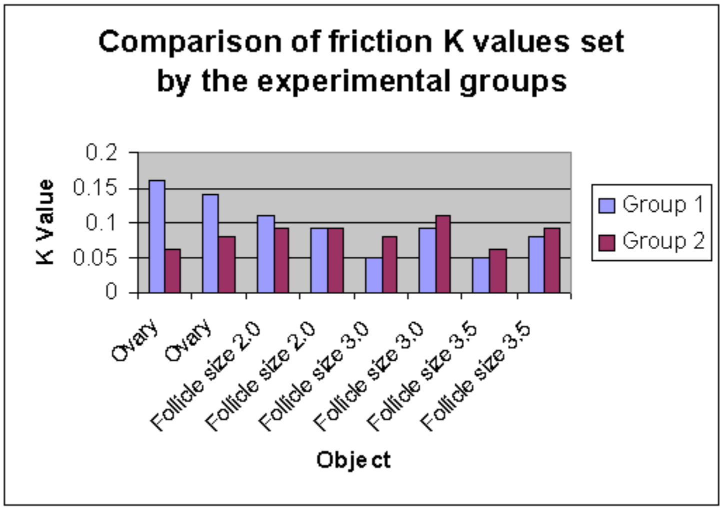


Figure 23 - Chart showing the comparison of friction k values for different objects between the experimental groups. Each object was presented twice.

The damping values for both groups are displayed in Figure 24. For damping, the values set were again far to the left of the slider for both groups. The only exception was Group 1 in setting damping for the ovaries. Again an extra expert was consulted. The higher damping value was judged to be closer to the feel of a real ovary and therefore the lower value was disregarded.

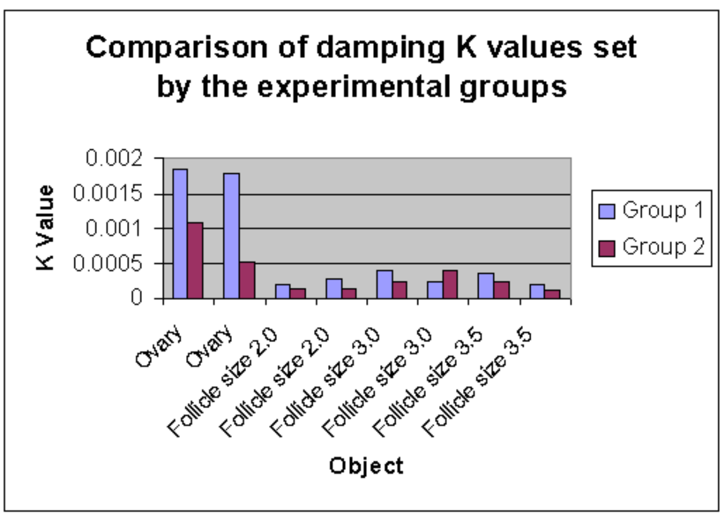


Figure 24 - Chart showing the comparison of damping k values for different objects between the experimental groups. Each object was presented twice.

Confidence levels for the models ranged between 3 and 4 out of 5 for Group 1. Group 2 were more cautious and return a confidence value of 3 in each case. This would indicate reasonable confidence that the models presented were an approximation to real life.

4.10.5 Observational Results & Conclusions

Except in the case of ovary stiffness, the slider bars for the surface properties were set far to the left of the range. This would suggest that the range of values could be reduced while increasing the granularity in subsequent experiments. Participants felt that although this was the case, small differences in the values were not noticeable, and a higher resolution would provide little benefit. It can be argued that individual differences in the haptic properties of real life horse ovaries would make small variances in the feel of the virtual models less important.

In equine ovaries, most of a follicle exists under the surface of the ovary. The experimental subjects also suggested moving the follicles further inside the ovary to increase the realism of the model.

Group 1 were more receptive to the use of the technology, which shows in their confidence levels for the task. During the task they repeatedly used phrases such as *"that's an ovary there"*, which indicates a level of recognition with the models is possible. The completed models were presented to an extra horse expert for confirmation of their similarity to real models. The expert consulted felt that the models were *"a good approximation to the real thing"*.

In all follicle cases, there was little difference between the values chosen by the experts. In the case of any conflict (such as in ovary stiffness), an extra expert was consulted in order to get a clearer idea of which setting was nearer to the real examination. While it is not possible to accurately represent contact with a real life ovary using the PHANToM, the veterinarians involved in the experiment were positive about the feel of the models and the potential of the technology for a palpation task. Through checking the final models with an expert, it was possible to determine that the models were a reasonable approximation. This stage of the development process was used to provide credibility for the simulator to other veterinarians.

4.11 Conclusions

This chapter has described work that has been carried out to develop models for integration into the Glasgow Veterinary Course. It has highlighted some of the problems that Glasgow University Veterinary School and others have in training for procedures of this kind where student opportunities are limited. The possibility of using a Virtual Reality simulator in conjunction with traditional training methods has been introduced, and the potential benefits have been discussed. This section is important when considering Question 1 from the key research question. It has shown that following a validated CAL design technique allows designers to discover problems with current training methods, and to visualise how and where a training simulator should be integrated into a course to provide benefit.

Further, an experiment has been discussed to obtain haptic property information for horse ovaries. Discussion after the experiment suggests that the models provide a reasonable approximation for the feel of a real ovary. The models have been developed to allow students to practise location and sizing of follicles on equine ovaries. As discussed in Chapter 3, evaluation of these models is now important. The next stage in the development process required the models to be validated. Chapter 5 describes the experiments carried out to test the effectiveness of the models.

5 Validation of the Simulator

5.1 Introduction

Question 2 from the key research questions is concerned with the training effects of the simulator. From Section 3.4, the importance of validation of a medical simulator has been established. Before a simulator can be used as a training tool, it must show that it provides some benefit to the users. Chapter 4 examined the feasibility of a training simulator, and a study was performed to provide ovary models that approximate the feel of horse ovaries. Experienced clinicians were involved to provide face validity to the simulator. This chapter describes three further studies that have been carried out to test the validity of the Glasgow Horse Ovary Palpation Simulator as a tool for training.

5.2 Comparison of Simulated Ovary Palpation Training Over Different Skill Levels

5.2.1 Background

This section describes an initial attempt to validate the Glasgow Horse Ovary Palpation Simulator by comparing performance of users of different skill levels. Experimental participants were asked to identify the position and size of follicles on the surface of virtual ovaries. The two experimental groups were made up of expert and novice users. Experienced large animal veterinarians were chosen as expert users, and second year veterinary students were chosen as novice users. Using Neufeld's definitions [85] for measurable factors of a training tool, this study was testing the construct validity of the simulator.

Studies have shown that simulator practice can be used to improve a user's performance on the simulator, as well as psychomotor skills, but there is little evidence to suggest that these improvements carry over to actual surgical procedures. One method that has been used to provide a validation of a training simulator is to compare the performance on the simulator of clinicians of different experience levels in the procedure being simulated. If the experienced group performs better than the novice group, it can then be assumed that performance on the simulator relies on skills that the experienced group are better at than the novice group. However, this in itself does not indicate that the using the simulator will provide improved performance in the real life task. If it was also demonstrated that repeated use of the simulator improved simulator performance, then it can be argued that the simulator is improving skills that will be useful for the real life task. This method therefore provides a good method for an initial validation of the models. It is commonly used as it provides a method of validation without the ethical issues involved in examining the effect of simulator training on real patients.

One example of such a study is presented by O'Toole *et al.* [89]. They describe a construct validity experiment that compares experienced surgeons' performance against medical students' on a surgical simulator. The simulated procedure involved inserting a curved needle into a vessel. Metrics were defined to measure performance in the task. These were: performance error score (an overall performance rating), tissue damage, surface damage, excess time, accuracy error, excess tool motion, angle error and peak force. There is some uncertainty about the metrics as no information was presented on how these metrics were developed or measured. Gorman [42] notes that in particular, metrics such as tissue and surface damage are difficult to develop. Twelve medical students and nine surgeons participated in the experiment. They demonstrated that the surgeons performed significantly better in the metrics performance error score, tissue damage, excess time, excess tool motion, and angle error. They conclude that their simulator may be useful in quantifying surgical skill.

Burdea *et al*. [16] attempted to validate a prostate simulator using a similar method. This experiment is described in detail in Section 3.4. The results showed that the non-medical students performed slightly better than the urology residents in the experiment although both were considerably lower in diagnosis accuracy than the control group using the rubber models. Burdea *et al.* conclude that improvements to the virtual models must be made to ensure that they are more similar to the real .

Sherman *et al.* [106] examine the performance of 29 trainee doctors of different levels of experience using the Virtual Environment Knee Arthroscopy Simulator (VE-KATS). This provides a graphical virtual environment for knee arthroscopy procedures, but does not provide any haptic feedback. The scoring system for performance on the simulator was based on a validated scoring system for arthroscopy procedures, as well as time to complete task. However, analysis of the results showed a poor correlation between simulator performance and trainee experience. Similar results were obtained when plotting time to complete task, and composite score (score/time to complete task) against trainee experience. One explanation given is that it is possible that more experienced surgeons rely on kinaesthetic cues from the tools, which are not present in the VE-KATS environment. The simulator presented in this thesis does provide haptic feedback so it is hoped that the problems Sherman *et al.* faced will not be experienced.

5.2.2 Method

5.2.2.1 Performance measures

This study will test the construct validity of the Glasgow Horse Ovary palpation simulator. The description of the ovary palpation procedure, the virtual models, and how they were developed has been discussed in Chapter 4. Participants in this experiment were asked to locate and size follicles on virtual ovaries using touch alone. As this is a diagnostic procedure the metrics chosen to provide information on the performance of the participants were very high level. The important questions in this procedure are where was the follicle found and what size was it. Time to complete task was also chosen as a metric as participants who are more experienced would be expected to be more comfortable with the task, and therefore complete it more quickly. Other factors mentioned above such as tissue damage would be applicable also. However, this is difficult to accurately calculate without first measuring the physical properties of the tissue. Workload results were also captured to determine how difficult the participant perceived the task to be.

For each examination, cursor position and force information were sampled every tenth time through the PHANToM servo loop. That is, the data was sampled at a rate of approximately one hundred samples per second. Samples at the rate of the PHANToM update rate of approximately one thousand times a second were not used, as the size of output files generated was restrictively large. These data were recorded to further analyse the similarities and differences in the performance of veterinary students and clinicians.

5.2.2.2 Training

As none of the participants had any previous experience in using the PHANToM, they were initially presented with the standard 'Blocks' demo developed by SensAble Technologies to familiarise them with the device. This demo allows a user to interact with two cubes in a three dimensional environment. It was used to allow the participants to familiarise themselves with using the PHANToM as a three dimensional input device. Participants could both feel their interactions through the PHANToM, and see their interactions on the screen.

The participants were then presented with an environment containing a haptic only representation of two spheres (shown in Figure 25). In this condition, they could not see the spheres or the PHANToM cursor on the screen.

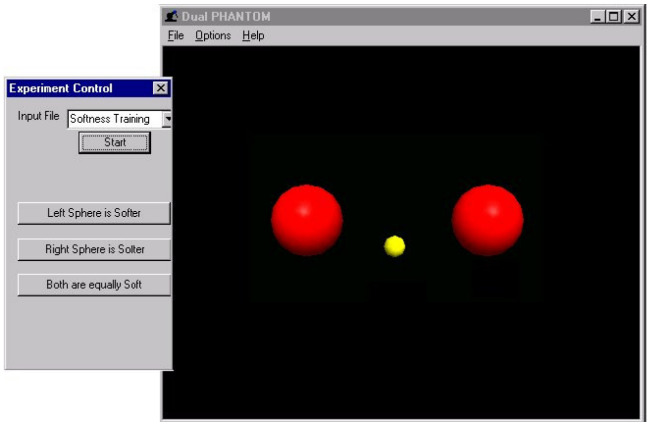


Figure 25. The virtual environment used to train participants in size and softness discrimination using the PHANToM. The participant's cursor is the smaller sphere in the centre. The participant received no visual feedback from the main window once the experiment had started.

Initially the size of the spheres varied. Each participant was asked to state whether the left or right sphere was larger or they were of equal size using touch alone. Three sizes of sphere were presented the participants. Each sphere was 1.5cm, 2.0cm, or 2.5cm in diameter. A random selection of each of these cases was presented to participants to explore until they reach an appropriate level of performance. There was no time limit for the exploration, but participants proceeded after answering five cases correctly. Participants answered a case by clicking one of three buttons in a dialog box using a mouse. The button labels were ‘Left sphere is larger’, ‘Right sphere is larger’, and 'They are the same size’.

A similar training session was provided for training in softness discrimination using the PHANToM. The same environment containing two spheres was used although the size of the spheres remained constant while their softness was varied. Each sphere used a linear force model that can be represented by the equation Reaction Force = constant (k) * penetration distance. To vary the softness, k was varied. Four k values were used in this training stage. They were 0.6, 0.45, 0.3, and 0.15. Participants stated whether the left or right sphere was softer, or whether they were of equal softness. Once again there was no time limit, and the participant proceeded after answering five cases correctly. Participants answered a case by clicking one of three buttons in a dialog box using a mouse. The button labels were 'Left sphere is softer', 'Right sphere is softer', and 'Both are equally soft'.

These training stages were designed to provide some initial familiarisation of the PHANToM as well as training in size and softness discrimination, and locating objects though touch alone. These skills were important for the experimental task.

In addition to the above training, users were presented with a visual and haptic representation of the HOPS environment (shown in Figure 32) before each experimental session. The participants were asked to explore the plain environment - no surface features - for five minutes. The time taken in completing the training varied between 18 and 25 minutes with a mean of approximately 21 minutes.

5.2.3 The Task

The experimental task involved identifying follicles on the surface of the virtual ovaries through touch alone. Participants were presented with the same eight cases but in a random order. Participants were aware that in each case there were zero, one or more follicles present on either ovary up to a maximum of five follicles in total. Each participant was given up to five minutes to explore the environment while identifying all follicles. Participants were made aware that they could finish the examination at any time if they felt they had completed the task before five minutes. Identification of a follicle involved identifying its position - either left or right ovary, front or back of the ovary, and top or bottom of the ovary - and its size. Participants were told that a follicle could be 2cm, 3cm or 3.5cm in diameter. Timing information for each case was calculated for analysis. Workload measurements were collected from all participants with a NASA Task Load indeX (TLX) workload evaluation form described by Moroney *et al.* [82] (see Appendix B.3).

The TLX consists of a series of six rating scales developed by NASA. The scales are unmarked, equal appearing interval scales with 20 values corresponding to the workload factors 'Mental Demand', 'Physical Demand', 'Time Pressure', 'Perceived Performance Level Achieved', 'Effort Expended', and 'Frustration Experienced'. In the original form of the TLX, a user would also use pairwise comparisons to rank each of the factors in terms of importance for the task and use this to give a unitary workload value for the task. Byers *et al.* [17] suggest that a simplified version of the TLX returns comparable results. Their studies suggest that by taking a mean value of the six workload factors, results comparable to the pairwise comparison method can be achieved. This thesis will use the method described by Byers *et al.* to calculate the workload. Similar scales to rank participant 'Confidence' and 'Fatigue' were also included on top of the TLX form given to the participants, however, data gathered from these scales were not used in the final workload calculation.

5.2.3.1 Participants

There were two subject groups involved in the experiment:

- Group A consisted of second year veterinary students from Glasgow University Veterinary School. At this stage in the course, students have some knowledge of horse ovary palpation through lectures, but have no practical ovary palpation experience.
- Group B consisted of experienced large animal veterinarians. Each participant had several years of experience and practice in large animal ovary palpation.

In this initial study, Group A contained 10 participants and Group B contained 7 participants.

5.2.3.2 Experimental Apparatus

During the experiment, users interacted with the virtual environments using a PHANToM 1.0 with the standard thimble attachment. The equipment was set up as shown in Figure 26 such that a participant received no visual feedback. Participants also wore headphones to mask any noise produced by the PHANToM motors.

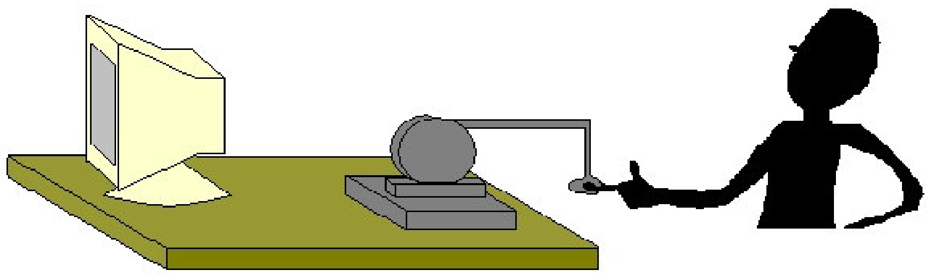


Figure 26. Experimental setup used for the task. The screen is turned away from the user so he or she receives no graphical feedback while feeling the ovaries.

5.2.4 Hypotheses

5.2.4.1 Hypothesis 1

The measured performance of the expert group in placing the virtual follicles will be significantly better than the performance of the novice group. This performance is measured as the number of correctly placed follicles. To correctly place a follicle, the participant must correctly state whether it is on the left or right ovary, the top or bottom of the ovary, and whether it is on the front or back of the ovary. The dependent variable was accuracy in placing follicles on the virtual ovaries. The independent variable was the experience level of the participant.

5.2.4.2 Hypothesis 2

The measured performance of expert group in placing and sizing the virtual follicles will be significantly better than the performance of the novice group. This performance is measured as the number of correctly placed and sized follicles. The dependent variable was accuracy in diagnosing follicles on the virtual ovaries. The independent variable was the experience level of the participant.

5.2.4.3 Hypothesis 3

The measured workload factors of the expert will be significantly lower than the workload factors of the novice group. Due to the familiarity of the expert group to the real life task, the required mental demand and effort will be less. The expert group will show a significantly higher confidence rating and display a higher perceived performance on the simulator. Workload is measure using a NASA TLX scale. The dependent variables were the individual workload factors. The independent variable was the experience level of the participant.

5.2.5 Results

All the data collected from the experiment is shown in Appendix B.4.

5.2.5.1 Performance on the Simulator

The results for performance in placing and sizing the follicles on the virtual ovaries are shown in Table 5. Each participant was exposed to the same 22 follicles. The mean number of follicles that were reported by the Novice Group was 18.8, compared to 18.4 follicles for the Expert Group. Of the 22 follicles the Novice Group correctly positioned a mean of 15.6, where as the Expert Group correctly positioned a mean of 13.6. These results were analysed using a Kruskal-Wallis [21]. This is a non-parametric statistical method to test for significance when two or more conditions are used. The data cannot be considered to be normally distributed as it is discrete constrained data and the sample size is small. The difference in this instance was found to be not significant (p = 0.46). The Novice Group correctly positioned and sized a mean of 8.0 of the 22 follicles, compared to a mean of 7.4 for Expert Group. These data were again tested using a Kruskal-Wallis test, and again the difference was shown to be not significant (p = 088).

	Mean Follicles found per trial	Mean Correctly Positioned	Mean Correctly Positioned & Sized
Novice Group	18.8	15.6	8.0
Expert Group	18.4	13.6	7.4

Table 5. Performance in identifying follicles on the ovary surface for both groups. Each participant was presented with the same 22 follicle cases.

The mean time taken for each group is shown in Figure 27. The Novice Group took a mean time of 242.5s for each examination, compared to 244.7s for the Expert Group. Analysis of these results using a GLM ANOVA [21] – the parametric equivalent of the Kruskal-Wallis test - showed the difference to be not significant (F_{16}=0.01, p=0.932).

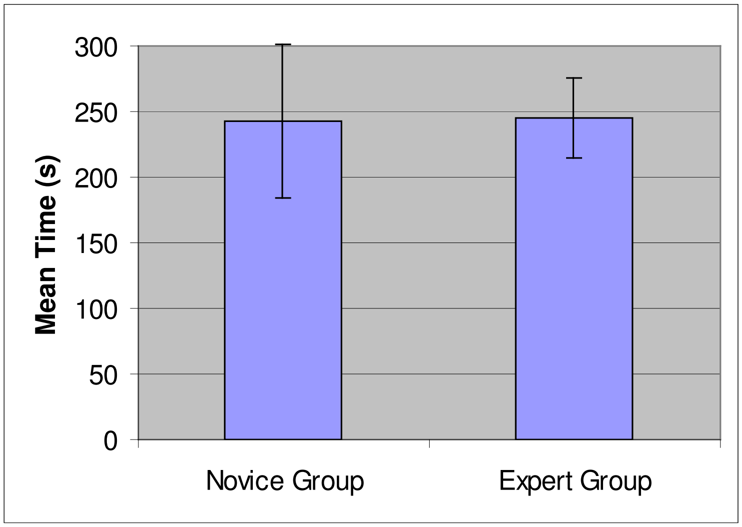


Figure 27. Mean time taken by each group to complete one examination. All examinations were capped at a maximum of 300s. Error bars indicate standard deviation for the data series.

Further analysis was conducted on the number of times a participant was stopped before his or her examination was completed at the maximum 300 seconds - referred to as timeouts - as this would affect the mean time taken for each examination. As each participant examined the same eight ovary cases, there was a maximum of eight timeouts per participant. The mean number of timeouts for the novice group was 3.8, compared to 2.7 timeouts for the expert group.

Finally, mean workload results for each group obtained using the NASA TLX scales are shown in Figure 28. It is important to note that the 'Performance Achieved' scale and the 'Confidence Level' have been inverted such that a lower score indicates a lower workload and therefore higher perceived performance or confidence in the task. Each of the factors was analysed individually using a Kruskal-Wallis test. There were no significant differences detected between any of the workload factors.

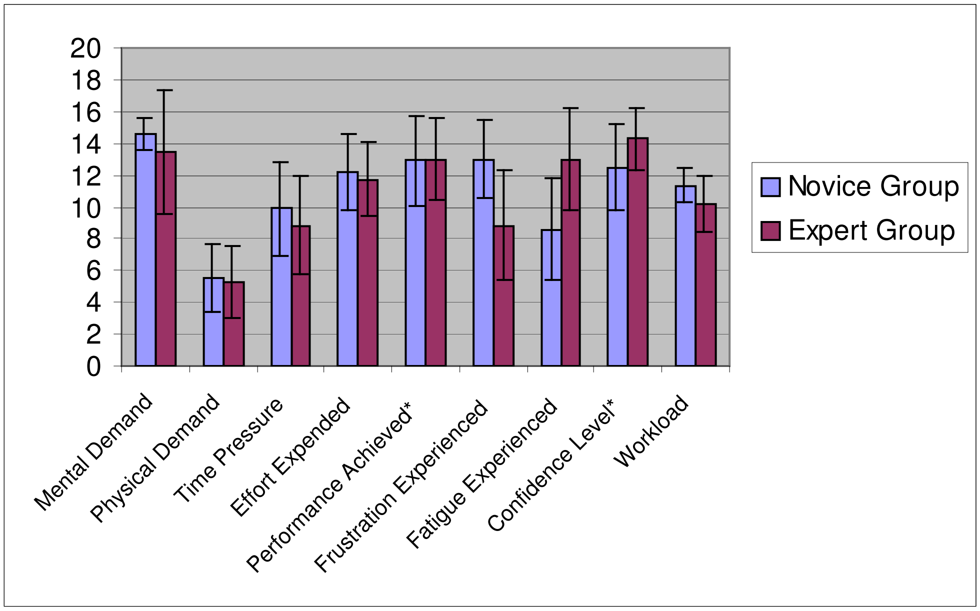


Figure 28. Mean workload results gathered from each group using the NASA TLX scales. A lower score indicates a lower required workload for the task. For 'Performance Achieved' and 'Confidence Level', a lower score indicates a lower workload and therefore higher perceived performance and confidence in the task. Workload is the mean of the six original TLX scales. Error bars indicate a confidence interval of 95% about the mean.

5.2.6 Discussion

5.2.6.1 Comparison of Placing and Sizing Follicles Data

Results suggest that Hypothesis 1 is not supported. Of the 22 follicles present in the experiment, both groups identified similar numbers of follicles. However, differences are noticed in the number of follicles correctly placed. The novice group placed a mean of 15.6 (70.9%) of follicles correctly, where as the expert group placed a mean of 13.6 (61.7%) correctly. Although this difference was not found to be significant when tested, it was expected that the expert group would correctly place follicles more frequently. This was confirmed when many of the expert group reported problems in placing a follicle on the ovary once found. Although palpation in a real examination often involves only one finger, the veterinarian will hold the follicle while palpating. He or she will therefore have an idea of the position of any follicle found with respect to the ovary in his or her hand. This is not the case in the virtual model since users are restricted to one point of contact with the environment. A user must trace the shape of the ovary to determine where he or she is on the ovary.

Differences were also observed in the techniques used by the novice and expert groups in the trials. Participants from the novice group tended to maintain contact with the ovary being search, and maintain a steady force to trace the shape of the ovary in a circular motion. Participants from the expert group tended to move across the ovary surface, repeatedly prodding it and therefore varying force. The initial method may be better for examining the surface of the virtual ovary, but this method of exploration is not possible in an actual examination. Another factor that must be considered is that as Berg *et al.* [7] note, the simulator may reflect differences in hand-eye coordination or gamesmanship. It is possible that the students treated the simulator as a game rather than a virtual model of a real life procedure.

5.2.6.2 Comparison of Timing Data

Hypothesis 2 has not been supported by the results. Although there are slight individual time differences when comparing between the two groups, it is interesting to note that on average, both groups took a similar length of time to complete the eight cases. The novice group took a mean of 242.5 seconds per examination compared to 244.7 seconds taken by the expert group. This equates to a mean difference of less than 20 seconds over the whole experiment that took both groups a mean of more than 30 minutes to complete. This difference is unsurprisingly not significant. The novice group did not complete on average 3.8 examinations out of the 8 cases within the maximum five minute period for each participant compared to 2.7 out of 8 for the expert group. If the examinations were not capped at this time, it would therefore be expected that the mean time for the participants in the novice group would be increased more than that of the expert group. It is unlikely however that this increase would lead to a significant difference between the groups.

5.2.6.3 Workload Analysis

Results would suggest that Hypothesis 3 has not been supported. There were no significant differences detected between any of the individual workload factors, or in the overall workload. It is important to note that the results obtained from the factors 'Performance Achieved' and 'Confidence Level' of the expert group suggest that the experts had a lower than expected confidence in the task and perceived performance. The mean values recorded from the expert group were 13.0 and 14.2 respectively. This could be due to the fact that locating follicles on an ovary is performed differently in the simulator and in the real examination. Placing the follicles on the ovaries with one contact point - without knowing when the ovary is in comparison to the examiners hand - presents a harder challenge than in the real examination.

The results gathered from the 'Mental Demand' factor suggest that both groups found the task very difficult to perform mentally. This is possibly due to the lack of experience in exploring an object with one point of contact rather than the whole hand. Further practice in using the PHANToM with no visual feedback before the task might reduce this value for the task.

5.2.6.4 Overall

There results suggest that it is not possible to distinguish between the different skill levels on the simulator. There are several possible reasons why this could be the case. Berg *et al.* [7] note that a comparison experiment may fail as it is testing one or more confounding factors rather than skill at the procedure. The expert group may not have been able to suspend their disbelief causing them to concentrate on the differences rather than the similarities to the real procedure. This group would have been trying to make the association between the virtual model and the real life object. The novice group were more likely to treat the experiment as a game, and not try to relate the features to the real life procedure. The models are designed to be approximation to the real life object.

5.2.7 Cursor Trace

For all examinations, cursor position and the reaction force from an object were sampled at a rate of once every ten times through the PHANToM update loop, which approximates to about 100 times a second. This information was later analysed to determine how the different groups of participants used the simulator to explore the virtual ovaries. Three individual factors were examined:

- The peak forces used during the examinations for each participant. As the PHANToM device cannot measure force applied by a user, reaction forces from an object were used as an approximation of the forces used;
- The division of time spent at feeling different levels of reaction force from the objects in the environment. The reaction forces were split into categories from no force to greater than 3 Newtons, with categories in between of 0.3 Netwons range each;
- The division of time spent on the individual objects in the environment. Categories used for the analysis were touching 'Nothing', 'Left Ovary', 'Right Ovary', 'Follicle', and 'Rectal Wall'.

The findings from these analyses are now discussed in the following sections.

5.2.7.1 Peak Force

Mean peak reaction force felt by a participant for the eight examinations was recorded for each participant. For the novice group, the mean peak force felt during an examination was 2.27N compared to 2.22 for the expert group. The difference was found to be not significant when analysed using a GLM ANOVA (F_{16}=0.04, p=0.85). When examining the same data for gender differences, the mean male peak force was 2.23 compared to 2.25 for female participants. Again this was found to be not significant (F_{16}=0.06, p=0.81). This suggested that participants used similar values for peak force throughout the examinations, and as the novice group peak forces were not significantly different from the expert group peak force, forces used by the novice group would be reasonable for a real life ovary examination.

5.2.7.2 Division of Time at Different Reaction Forces

Figure 29 shows the mean number of samples spent at different reaction forces for each participant in the expert and novice groups. Again each sample represents approximately 0.01s of actual time. For both groups, the majority of time was spent during the examination at a reaction force of below 1 Newton. From Figure 29, it can be seen that both groups spent similar amounts of time in each of the different force categories. From the graph, the expert groups appears to have spent longer per participant in the 0.3N to 0.6N category than the Novice group. This difference proved not to be significant when analysed using a Kruskal-Wallis test (p=0.38).

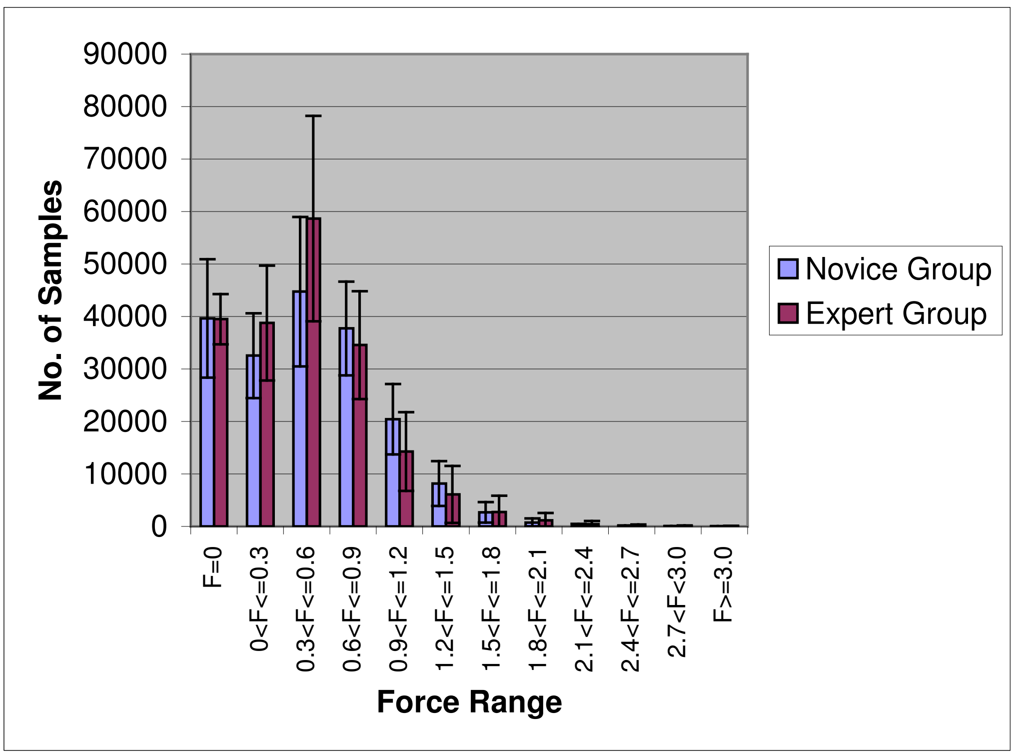


Figure 29. The mean number of samples for each group in the different force ranges. Error bars indicate a confidence interval of 95% about the mean.

5.2.7.3 Division of Time Spent on Each Virtual Object

Figure 30 shows mean time spent by participants in each group touching the different virtual objects within the scene. There are again considerable similarities between the profile of the novice and expert group and this is borne out in the statistics as there are no significant differences between the number of samples detected on any of the virtual objects. It is interesting to note that for both groups the longest time spent in contact with any one object was the follicles. The follicles were the smallest objects in the scene but were touched for the longest amount of time during an examination. This could be because of the time required to size a follicle once located. The results suggest that participants found sizing the virtual follicles difficult as less than half of the follicles during the experiment were found and sized correctly.

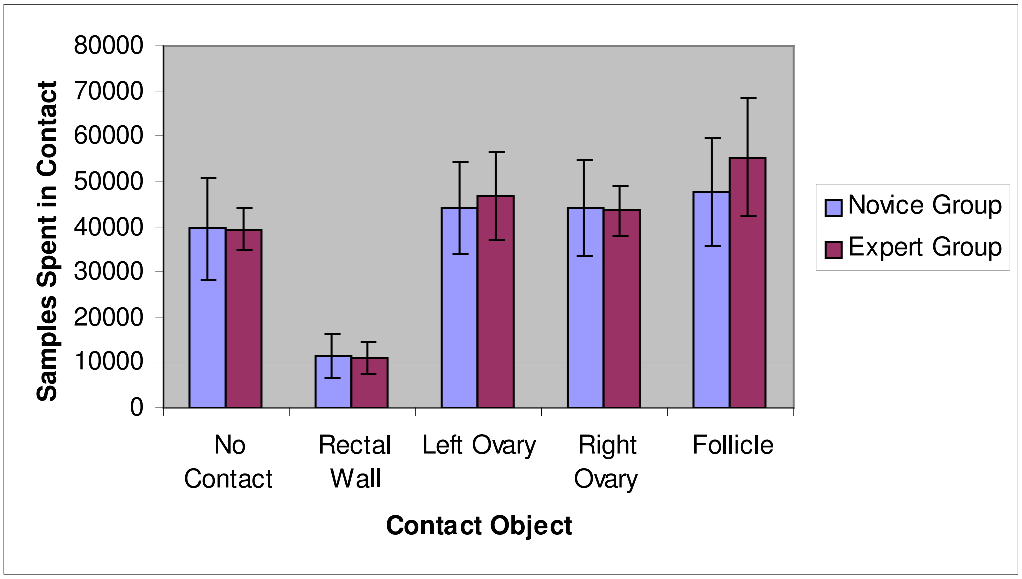


Figure 30. The mean number of samples spent in contact with different scene objects. The error bars indicate a confidence interval of 95% about the mean.

5.2.8 Conclusions

These results indicate that the simulator was not able to distinguish between the performance of novices and experts. They indicate that the experienced horse veterinarians found examining the virtual ovaries more difficult than was expected and the novice group performed better on the simulator than expected. During the experiment, the cursor path and reaction force information collected proved not to be significantly different between the novice and expert groups for the measured factors. Some possible reasons for these findings have been discussed.

From the results of this experiment, we were not able to validate the simulator. It was therefore important to examine other forms of validation. One limitation of testing the construct validity of a simulator is that it does not examine the training effects of the simulator. A longer term study would allow analysis of any improvements in performance over several sessions on the simulator. Comparison of students trained on the simulator with students trained in the traditional manner could then be used to evaluate the benefits of simulated ovary palpation training.

5.3 A Study to Evaluate Training Effects of the HOPS Simulator

5.3.1 Introduction

There have been previous studies that examine simulator performance over a number of training sessions. Gorman *et al.* [42] present a study examining the performance on a simulator involving driving a simulated needle through a blood vessel. Their study noted a significant improvement over three training sessions, and they conclude that the simulator may be used to improve the psychomotor skills required for surgery. They note however, that there is no evidence that the improvement in skills demonstrated on the simulator would carry over to the real task.

Kothari *et al.* [60] present a similar study in which five training session - using the MIST-VR simulator - are performed by 13 participants over a period of five consecutive days. No data were presented that show improvements or otherwise in performance of the participants on the simulator over this time. However, improvement was shown in using minimally invasive surgical tools to tie a knot in a piece of foam. They note that this may also be due to the fact that participants performed the procedure several times before the training in order to provide a baseline for performance measures.

This study takes an in depth look at performance changes due to repeated usage of a palpation simulator. The metrics task performance, timing, and perceived workload were recorded and analysed over multiple equally spaced training sessions. Retention of skills was also tested through a further training session after a longer break from the simulator. Unlike the previous studies mentioned, the focus was not on measuring improvements in psychomotor learning. Participants were rated on their abilities to distinguish shape, compliance and size data for the virtual objects through touch alone. These are skills required in a real large animal ovary examination.

This is the initial stage of testing the concurrent validation of HOPS. A further experiment then compared the performance of students trained using the simulator to those that have received traditional training from the Glasgow University Veterinary School.

5.3.1.1 Participants

One group of participants was involved in the experiment. The group consisted of second year veterinary students at Glasgow University Veterinary School who had learned the theory of ovary palpation from lectures and labs, but had no practical experience of in-vivo ovary palpation through university teaching. There were eight participants with seven female and one male. All were regular computer users, but none had any previous exposure to the PHANToM.

5.3.2 Hypotheses

5.3.2.1 Hypothesis 1

The performance level of the participants on the simulator will significantly improve with an increasing number of training sessions. Performance is defined by the number of follicles that have been placed correctly on an ovary, and the correct sizing of these follicles. The dependent variable for Hypothesis 1 is the performance level on the virtual ovaries. The independent variable is the quantity of haptic training that the participants receive.

5.3.2.2 Hypothesis 2

The time required to complete the task will significantly decrease as the participant receives more simulator training. The dependent variable for Hypothesis 2 is the time taken to complete the examination of the ovaries. The independent variable is the quantity of haptic training that the participants receive.

5.3.2.3 Hypothesis 3

The measured individual workload factors when examining the virtual ovaries will significantly decrease as the participants receive more simulator training. Participants' overall workload will significantly decrease while confidence will significantly increase. The dependent variables for Hypothesis 3 are individual workload, confidence and overall workload when examining the virtual ovaries. The independent variable is the quantity of haptic training that the participants receive.

5.3.2.4 Hypothesis 4

The performance on the virtual ovaries will not significantly decrease when a longer amount of time is left between training sessions. Performance is measured as in Hypotheses 1 and 2. The dependent variable for Hypothesis 4 is the performance level on the virtual ovaries. The independent variable is the time between haptic training sessions.

5.3.3 Experimental Set-up

Participants interacted with the HOPS environment using a PHANToM 1.0 from SensAble Technologies with the standard thimble attachment. The simulation was run on a 700 MHz dual-processor Windows PC. The equipment for the experiment was set up as shown in Figure 31. The participants wore headphones to mask noises produced by the PHANToM motors.

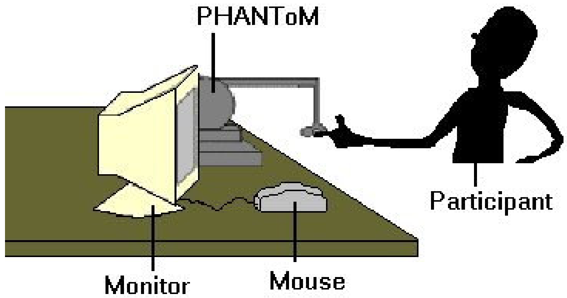


Figure 31. The experimental setup used. Participants interacted through the PHANToM. The virtual models and PHANToM interactions were not displayed to the user on-screen, while the mouse was used to start and stop examinations

5.3.4 Training

The training for this experiment was identical to the training given in the previous study. All participants completed the training successfully although no timing information was taken in this case. This training session was only presented to the user before the first experimental session.

5.3.5 Task

The HOPS environment models in this experiment were designed in close collaboration with experience veterinarians. The method for building the models is described in Chapter 4.

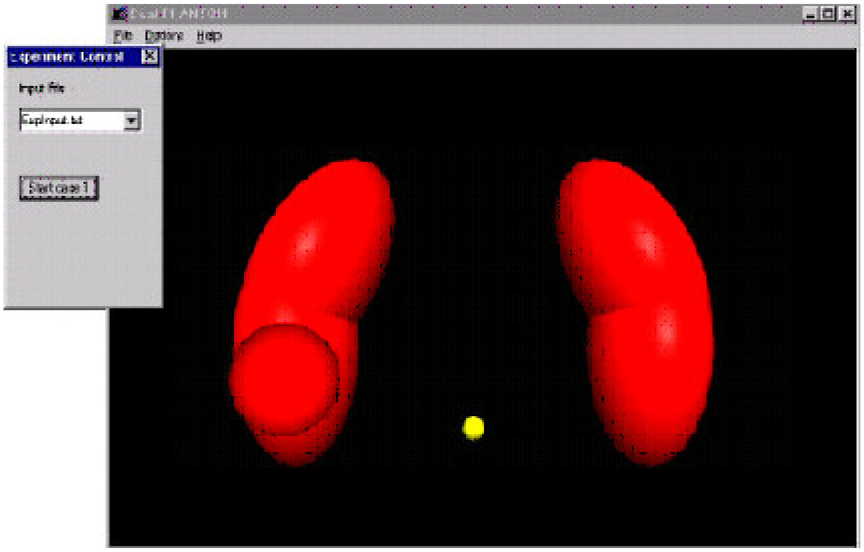


Figure 32. The Horse Ovary Palpation Simulator. This environment consists of a left and right ovary. On the bottom half of the left ovary, a spherical follicle can be seen. The user's cursor is shown as the sphere in the centre. The dialog box button on the left is used to start and stop cases.

The study took the form of a within groups repeated measure experimental design. All participants were presented with the same thirty-two ovary cases over four training session spaced a week apart, but in counterbalanced orders. Of these cases, there were twelve three follicle cases, twelve two follicle cases and eight one follicle cases. There were four orders of presentation with two participants being presented with each ordering. In each experimental session, participants were each presented with eight ovary cases. Also, in each experimental session, the total number of follicles in all cases was kept constant at seventeen. For each case, participants were told that there were zero, one or more follicles present on either ovary and were allowed five minutes to explore the environment while identifying all follicles. Identification involved positioning a follicle on the left or right ovary, the front or back of the ovary, and top or bottom half of the ovary. Once identified, participants were asked the size of the follicle. They were told that the follicles would either be 2cm, 3cm or 3.5cm in diameter. Participants started and stopped each case using the mouse. If a case was explored for five minutes, he/she was informed that the time for examining the current case was finished and was allowed to proceed to the next case. Participants provided answers for each case by filling in a results sheet. The section of this sheet relating to one ovary case is shown in Figure 33. The full results sheet is shown in Appendix C.

Follicle	Position			Size(cm)
	Ovary	Side	Height	
1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) **1 2 3 4 5** (High)

Figure 33. This form was filled in by each participant for each ovary case. He/she was asked to provide position information in the Ovary/Side/Height boxes and size information for each follicle. He/She also indicated the confidence that the diagnosis was correct.

Time measurements were taken for each case. As timing information would be affected by the number of follicles found in a case, there were equal numbers of cases of equal complexity in each experimental session. Therefore, there were two one-follicle cases, three two-follicle cases and three three-follicle cases presented in each session.

Participants were asked to complete a NASA TLX workload evaluation form after each experimental session had ended. The workload form was identical to that of the previous experiment and therefore contained the additional scales **'**Confidence**'** and **'**Fatigue**'.** Four such experimental sessions spaced a week apart were performed by each participant.

Participants took part in a fifth training session one month after the fourth session. They were presented with the same ovary conditions as they were in the first session. During this time, participants did not have any courses on reproduction, or perform any ovary examinations.

Participants were not told if their answers were correct or incorrect at any time during the experiment. This was to ensure that measured workload values were not affected by results, and that all training was as a result of time spent using the simulator.

5.3.6 Results

5.3.6.1 Performance on the Simulator

Correctness data for positioning the follicles is shown in Figure 34. Results were analysed using a Kruskal-Wallis test [21] using the number of training sessions as a factor. Only results from the first four training sessions were included as the fifth training session was only used to test Hypothesis 4. Increasing the number of training sessions was not found to have a significant effect when comparing mean accuracy in placing follicles on the ovaries over the four training sessions ($p = 0.065$). However, this may be due to the fact that the mean values quickly come close to the maximum possible value, with over 15 out of 17 follicles correctly placed during the third training session (as can be seen from Figure 34).

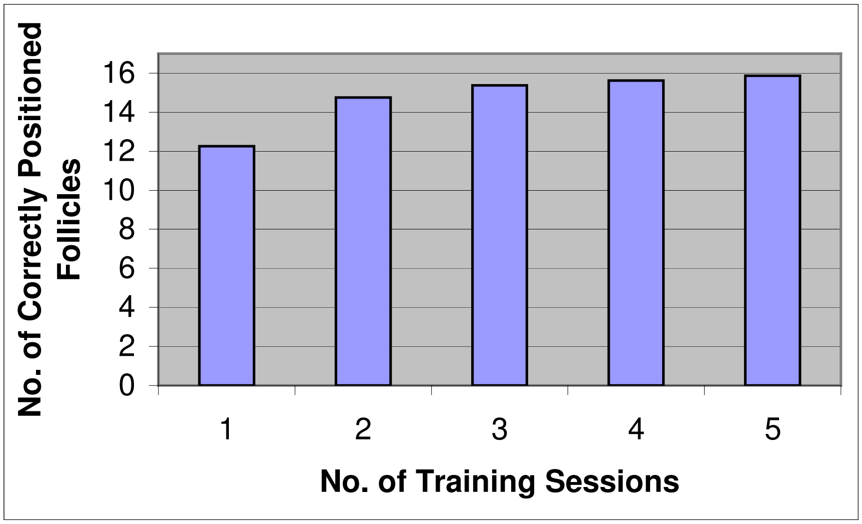


Figure 34. Mean number of correctly positioned follicles for all participants over 5 training sessions. There were 17 follicles for each participant in each session.

A Mann-Whitney test – a non-parametric test for significance between two samples - shows that there is also no significant difference between performance of participants in placing follicles between training sessions 4 and 5 ($p = 0.69$).

Similar analysis was carried out on follicles that were correctly positioned and sized over the four training sessions. The results of all five training sessions are shown in Figure 35. Although this data is discrete and constrained data, a normal distribution assumption will be made here for clarity. This is possible because in this instance, the means of data are in the middle of the upper and lower bounds and can therefore be considered to be unaffected by the constraints. The advantage of ANOVA analysis in this case is that it allows *post-hoc* Tukey analysis to be carried out on the data. General Linear Model ANOVA [21] analysis shows a significant performance difference as training progresses through the first four training sessions ($F_{3,21}=7.28$, $p<0.021$). *Post-hoc* analysis using a Tukey HSD test [21] revealed that there were significant differences between performance in sessions 1 and 3 ($p<0.003$), sessions 1 and 4 ($p<0.02$) and sessions 2 and 3 ($p<0.04$). Although a slight decrease in performance can be seen in Figure 35 between sessions 3 and 4, this difference is not significant.

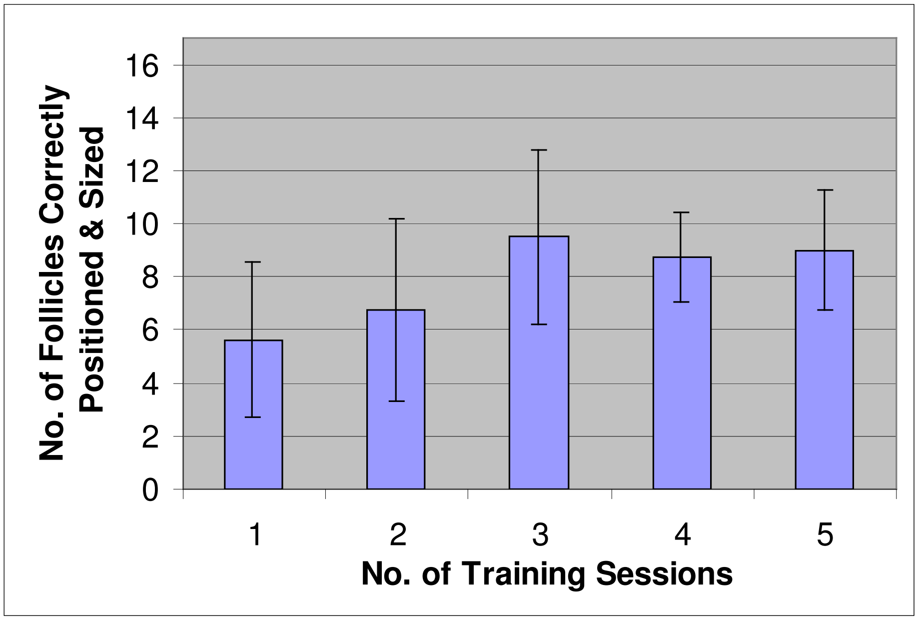


Figure 35. Mean number of correctly positioned and sized follicles for each participant over five training sessions. There were 17 follicles for each participant in each session.

A GLM ANOVA shows that there is also no significant difference between performance in correctly placing and sizing follicles of participants between training sessions 4 and 5 ($F_{1,7} = 0.1$, $P = 0.76$).

5.3.6.2 Time Taken to Complete an Examination

The results of the timing data are shown in Figure 36. Timing data was again analysed using an ANOVA test. The results show a significant decrease in time taken to complete the task as training progressed ($F_{3,21}=10.64$, $p<0.001$). A *post-hoc* Tukey HSD test revealed a significant decrease in time taken for the task during sessions 1 and 2 ($p<0.05$), 1 and 3 ($p<0.03$), 1 and 4 ($p<0.001$), and 2 and 4 ($p<0.05$). Again, although time taken to complete the task decreases between sessions 3 and 4, this difference is not significant.

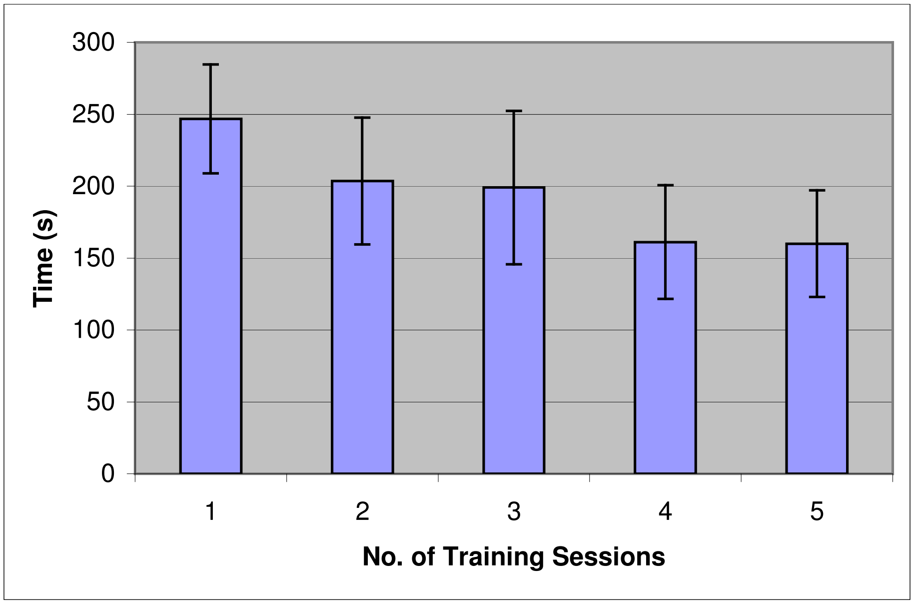


Figure 36. Mean time taken to examine one ovary case measured for all participants over 5 training sessions. The time for one examination was capped at 300s. Error bars indicate the standard deviation about the mean.

Table 6 shows the mean number of examinations stopped by the experimenter at 300 seconds for each participant over the 5 training sessions. It shows a consistent drop in the number of examinations that would have taken longer than the maximum time to complete.

Training Session	1	2	3	4	5
Mean examinations stopped at 300s	3.25	1.5	0.88	0.13	0

Table 6. The mean number of examinations for each participant that were stopped by the experimenter over the five training sessions. Each participant examined eight ovary cases per session.

A GLM ANOVA shows that there is no significant difference between time taken for the task between training sessions 4 and 5 ($F_{1,7} = 0.01$, $p = 0.94$).

5.3.6.3 Subjective Measures

The results gathered from workload analysis are shown in Figure 37. It is important to note that for 'Performance Achieved' and 'Confidence Level', a lower score indicates better perceived performance or confidence.

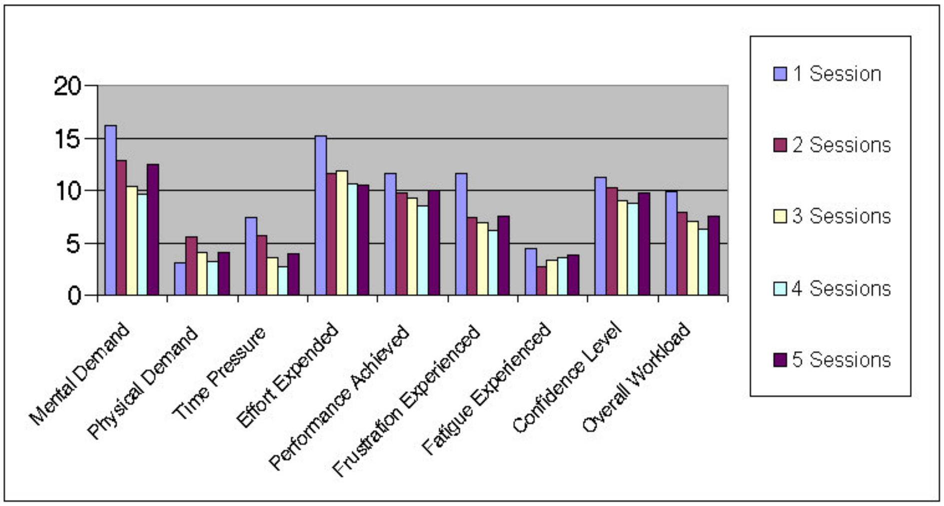


Figure 37. Mean workload for all Participants shown over 5 training sessions. 'Performance Achieved' and 'Confidence Level' are such that a lower value indicates better perceived performance or confidence.

Statistical analysis using a Krukal-Wallis test - with number of training sessions as the factor - suggests that there is no significant decrease in overall workload due to participants completing multiple training session (p=0.12). Similar tests (results shown in Table 7) on the individual factors suggest that the only workload factor significantly affected by multiple sessions is 'Time Pressure' ($p<0.05$). 'Mental Demand' also tends towards a significant value (P = 0.059).

A Mann-Whitney test was used to look for differences in workload factors between training sessions 4 and 5. Each of the workload factors was analysed individually, revealing no significant differences between these factors. Important for Hypothesis 4, confidence was not shown to have been affected by the one month gap between sessions (p = 0.91). Overall workload was also not significantly different between sessions (p = 0.56).

Workload Factor	1 TS Mean	2 TS Mean	3 TS Mean	4 TS Mean	P
Mental Demand	16.1	12.9	10.4	9.6	0.059
Physical Demand	3.1	5.5	4.1	3.3	0.62
Time Pressure	**7.4**	**5.6**	**3.6**	**2.8**	**0.048**
Effort Expended	15.1	11.6	11.9	10.6	0.46
Performance Achieved	11.6	9.8	9.3	8.5	0.45
Frustration Experienced	11.6	7.4	6.9	6.1	0.19
Fatigue Experienced	4.5	2.8	3.4	3.6	0.91
Confidence	11.3	10.3	9	8.8	0.76
Overall Workload	10.3	8.8	7.7	6.8	0.12

Table 7. The results of a Kruskal-Wallis analysis over workload data using number of training sessions as the factor. Significant factors are shown in bold.

5.3.7 Discussion of the Virtual Training Results

Hypothesis 1 has been supported by the results. As the number of experimental sessions increased, a significant increase in performance was also noted when examining correctly positioned and sized follicles. For correctly positioned follicles, participants quickly reached a high level of performance scoring a mean of 15 or more follicles correctly positioned from the third training session onwards. Results suggest that sizing the follicles proved a more difficult task that required more training. Significant performance differences were noted between Session 1 and all other Sessions as well as between Sessions 2 and 3. This suggests that participants were still learning, and improving performance during Session 3. There is a slight decrease in performance between Sessions 3 and 4, but this is not significant. This may indicate that the performance improvements due to time spent on the simulator had levelled off by Session 4.

Hypothesis 2 has also been supported by the results. Significant decreases in time were shown between Session 1 and all other Sessions, and Session 2 and 4. This difference is emphasised by the fact that examination times were capped at five minutes for each case. The majority of examinations not completed in five minutes occurred in the first week. Again, there was no significant time difference noted between Sessions 3 and 4. This may suggest that participants would not become much faster with more training. The timing data combined with the performance data shows that as participants received more training, they were able to complete the task in less time without having a detrimental effect on performance.

Hypothesis 3 was not supported by the data. There was no significant overall workload difference noted throughout the four training sessions. Although no significant decrease was noted, examination of the graph of the workload results suggests downward trend may be present in the mean overall workload data. Similarly, although there was a general increasing trend in mean confidence noted throughout the training sessions, these increases are not significant. The increase suggests however that a significant result may be possible with more training sessions. For each of the individual workload factors, only time pressure showed a significant decrease throughout the training sessions. This may correspond to the fact that the participants were performing better at this stage, and taking less time to complete an examination, and therefore finding it easier to complete the examination within the time limit. The mental demand factor also tended towards significance, which may suggest that participants became more comfortable with the task and found it easier to complete within the 5 minutes provided as the number of training sessions increased. This is also supported by the decrease in time taken for the task throughout the training sessions as noted above.

There is some data that suggests Hypothesis 4 may be supported by the results. No significant performance differences found between sessions 4 and 5 despite a one month interval between sessions. It is important to note that in both sessions, participants managed similar mean scores for placing and sizing the follicles. Similarly with time taken to complete the task, mean time taken was similar in both training sessions. Participants achieved similar levels of performance between the two sessions while taking a similar time to complete the task. Workload factors emphasise the fact that users found the task a similar level of difficulty. No significant differences were found between any of the workload factors. The participants' confidence was not shown to significantly decrease in the time period, and overall workload ratings remained similar.

These results provided an encouraging basis for future experiments. They demonstrated that learning takes place when using the HOPS simulator environment over several training sessions. The next stage of the experiment will demonstrate whether or not these improvements carry through to real life palpation skills.

5.4 A Specimen Ovary Examination Study

The results described above were the first stage in an experiment to assess the usefulness of HOPS as a training tool for veterinary students. The results show that as participants received more training on the simulator, their performance in diagnosing the condition of the ovaries improved, while time taken to make the diagnosis and mental demand for the task decreased. The second stage of the experiment was designed to measure how closely these improvements translated to improvement in performance in a real life ovary examination. This involved examining the performance in ovary palpation of participants trained using the HOPS simulator, and participants trained using traditional methods.

From Chapter 4, it can be seen that students at Glasgow University Veterinary School are introduced to large animal reproductive systems through exposure to specimen cow tracts. As these are used in tradition training methods, it can be assumed that they provide the most feasible method for testing the concurrent validity of HOPS.

5.4.1 Participants

There were two groups of participants in this study:

- **Group A** - eight participants trained using the HOPS simulator in the section above.
- **Group B** - eight participants trained using traditional methods.

Group A consisted of second year veterinary students who participated in the multi-session virtual training experiment described before. Each participant had received five training sessions on the HOPS simulator, but had no previous experience of real cow or horse ovary palpation. Group B consisted of second year veterinary students from Glasgow University who had no exposure to virtual training, but had participated in one two hour anatomy lab examining the bovine reproductive system as well as a reproduction lecture course. This anatomy lab is a standard section of the second year course.

5.4.2 Hypotheses

5.4.2.1 Hypothesis 1

Group A will perform significantly better on the specimen ovaries than Group B. Performance is based on the correct location and sizing of follicles. The dependent variable is the performance level on the specimen ovaries. The independent variable is the type of training: either haptic training or traditional training. Although both groups received theory from lectures, only the traditionally trained group received a two hour practical lab.

5.4.2.2 Hypothesis 2

Group A will take significantly less time than Group B when examining the specimen ovaries. The dependent variable is the time taken to examine the specimen ovaries. The independent variable is the type of training: either haptic training or traditional training.

5.4.2.3 Hypothesis 3

Group A and Group B, the measured overall workload on the specimen ovaries of Group A would be significantly lower than the overall workload of Group B. Group A will show a significantly higher confidence rating than Group B. The dependent variables are the individual and overall workload factors and confidence when examining the specimen ovaries. The independent variable is the type of training: either haptic training or traditional training.

5.4.3 Experimental Set-up

During the experiment, specimen tracts were obscured from the participants using a barrier. The barrier had a curtained opening to allow participants to feel the ovaries on the other side of the barrier without being able to see them, as shown in Figure 38. The ovary cases were each placed on different trays but with the same alignment such that for different cases, a participant would feel the different objects in the tracts in similar relative positions. Participants were asked to locate and size follicles on the ovaries of the specimen tracts provided.

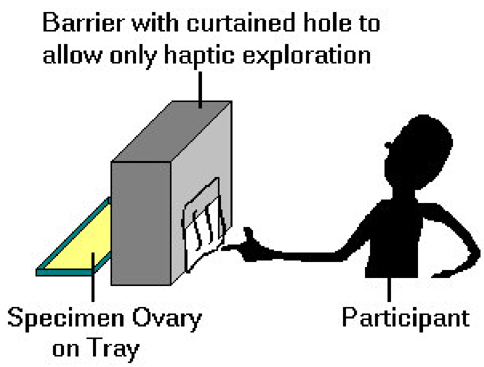


Figure 38. Experimental set-up for the specimen ovary examination

5.4.4 Design

This experiment used a between groups design.

5.4.4.1 Initial Introduction

Both groups of participants were given the same initial introduction immediately before the task. They were briefly shown a specimen tract - that would not be used in the experiment - and the uterus, uterine horns, and ovaries were identified for them. Although they could see the structure, they were not allowed to feel the tract. This training section lasted no more than thirty seconds in all cases.

5.4.4.2 Specimen Examination

Participants were presented with eight bovine specimen ovary cases. The experimental set-up shown in Figure 38 prevents participants from seeing the specimens, so they were restricted to haptic exploration only. Participants were further restricted to haptic exploration with one hand, as this restriction is also true for ovary examination in a cow or horse. Although there were eight cases presented to the participants, there were only four specimen ovaries available. Each specimen was also presented to the participant reversed such that the left ovary became the right ovary and *vice versa*. The participants were not made aware of this. Each specimen was presented for a maximum of five minutes, although a participant could stop the examination at any time before the five minutes when he or she felt that the examination had finished.

Participants were asked to identify the position (either left or right ovary), and diameter of all follicles on each specimen. It is important to note however that other structures such as corpus lutea existed on the ovaries as well. Participants therefore had to make the distinction between follicles and these other surface features that may feel similar to a novice. When a participant placed a follicle on a specific ovary, the participants answer was matched to the nearest unmatched follicle in size on that ovary. In the case where no follicle existed on the ovary, the participants answer was matched to another surface feature if possible. If no surface feature existed on the ovary that could be matched to the answer, this was noted. Time for task completion data was recorded in each case using a stopwatch. Workload data was captured using a NASA TLX workload evaluation form as before.

5.4.5 Results

5.4.5.1 Performance on the specimens

Correctness data for positioning the follicles on the specimen ovaries is shown in Figure 39. Results were analysed using a Mann-Whitney test. Mean scores for placing follicles for Group A and Group B were 12.9 and 11.8, respectively. This difference between the two groups in the accuracy of placing follicles on the specimen ovaries was not found to be significant ($W_{14} = 81$, $p = 0.16$).

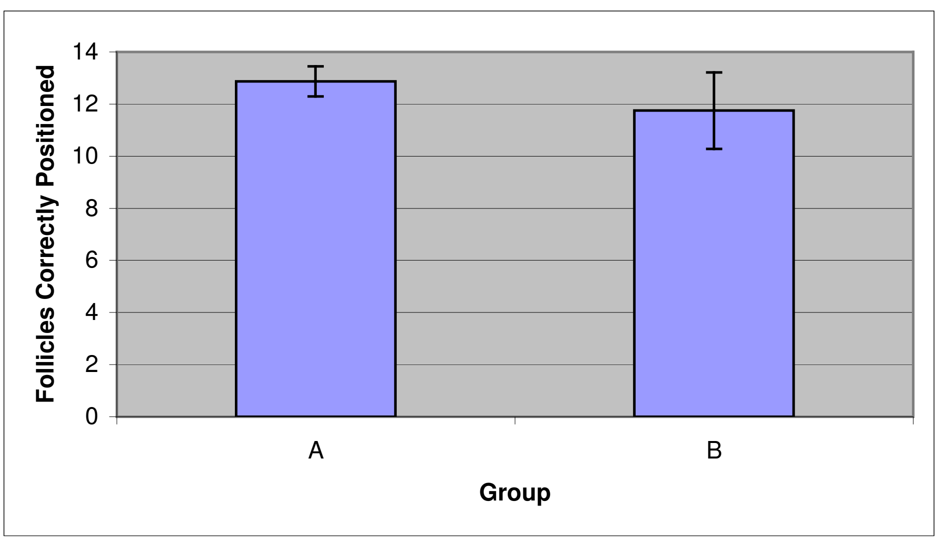


Figure 39. Mean number of follicles for each participant that were correctly positioned on the ovaries (out of 14). Error bars show a confidence internal of 95% about the mean.

The next factor analysed was the number of correctly placed follicles that were also estimated within 0.5cm of the actual diameter. This is within the accuracy expected to be achieved by the student. Size results for one Group A participant had to be discarded as his or her answers were not in the correct format. He or she used small, medium, and large to refer to sizes as opposed to specifying a size in centimetres. Again using a Mann-Whitney test, no significant difference was found between groups ($W_{13} = 66.5$, $p = 0.24$). Group A had a mean score of 9.6 follicles, where as Group B had a mean score of 7.5. The results are shown in Figure 40.

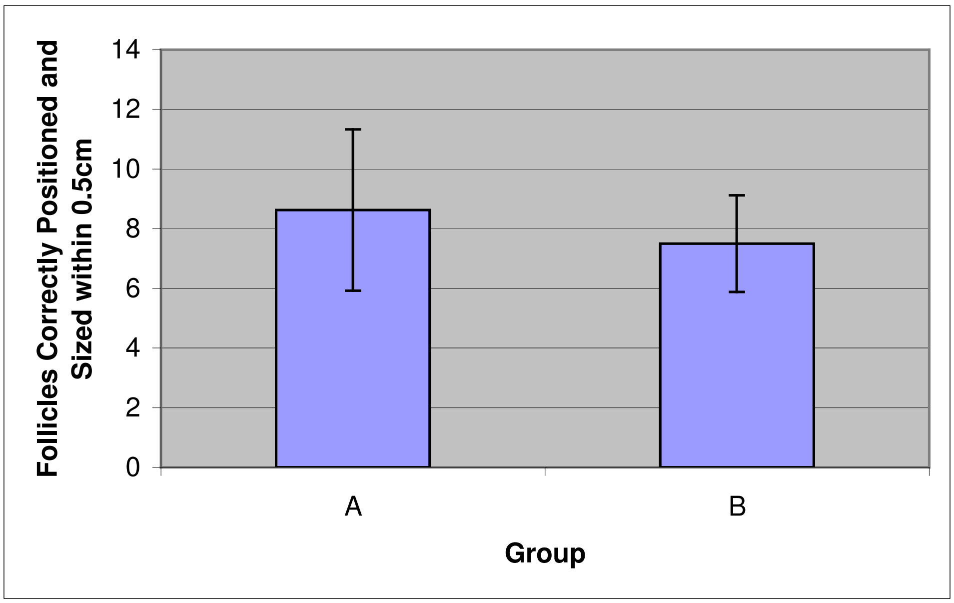


Figure 40. Mean number of follicles for each participant that were correctly positioned and sized within 0.5cm of the actual size (out of 14). Error bars show a confidence interval of 95% about the mean.

The mean distance from the correct size of a follicle was also analysed, this time using a two sample T-test. Again size results for one Group A participant had to be discarded. For Group A, the mean distance from the correct size was 0.48cm compared to 0.60cm from Group B. Again, this difference was not significant ($T_{13} = 0.79$, $p = 0.45$).

The above results all examine the correctness of the responses given by participants. Errors made during the study were also examined. No difference was found when looking at other surface features mistaken for follicles ($W_{14} = 78.0$, $p = 0.31$). The mean number of surface features mistaken for follicles was 6.25 for Group A and 4.88 for Group B. Extra follicles added where no follicle or other surface features exist were also analysed. Group A had a mean score of 4.13 compared to 3.63. This difference is not significant ($W_{14} = 66.5$, $p = 0.91$).

The final errors analysis, was the total number of errors made. These comprised of missed follicles, surface features mistaken for follicles, and extra objects that could be neither a follicle or other surface feature. Group A made a mean of 11.5 errors, where as Group B made a mean of 10.63. The difference between these values is not significant ($W_{14} = 66.0$, $p = 0.87$).

5.4.5.2 Time taken to complete an examination

A two-sample T-test was also used to examine differences between groups in time taken for an examination. The results are shown in Figure 41. Mean time taken by Group A was 200.5 seconds, where as mean time taken for Group B was 225.4 seconds. This difference is again not significant ($T_{14} = 1.40$, $p = 0.31$). An analysis of the number of timed out examinations - examinations that were stopped by the experimenter after five minutes - showed that the mean for timed out examinations for Group B was 2 out of 8 compared to 0.75 out of 8 for Group A.

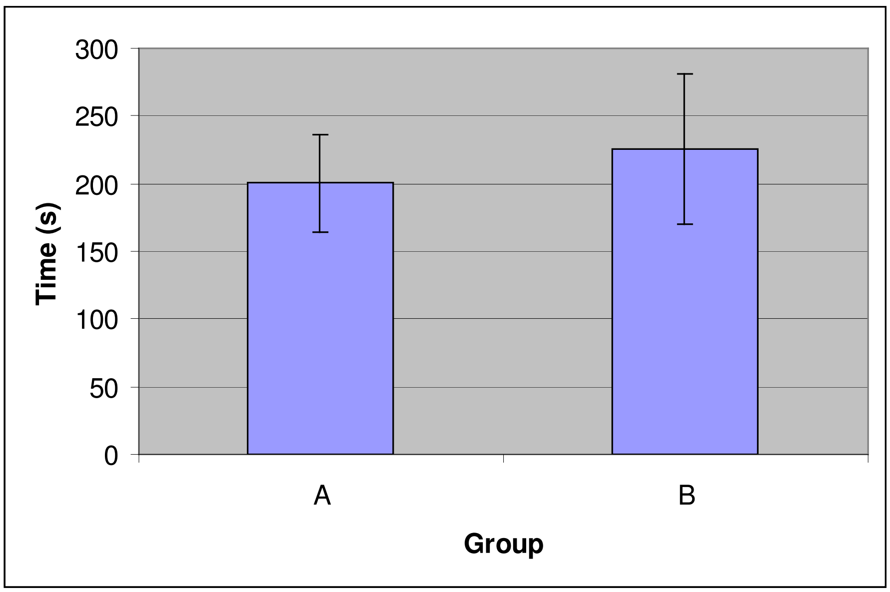


Figure 41. Mean time to complete one examination for either group (maximum of 300s). Error bars show the standard deviation about the mean.

5.4.5.3 Subjective measures

The results of the workload analysis for each group are shown in Figure 42. Each of the different workload elements was analysed individually, as well as the overall workload, which consists of the mean score for mental demand, physical demand, effort expended, time pressure, perceived performance achieved, and frustration experienced.

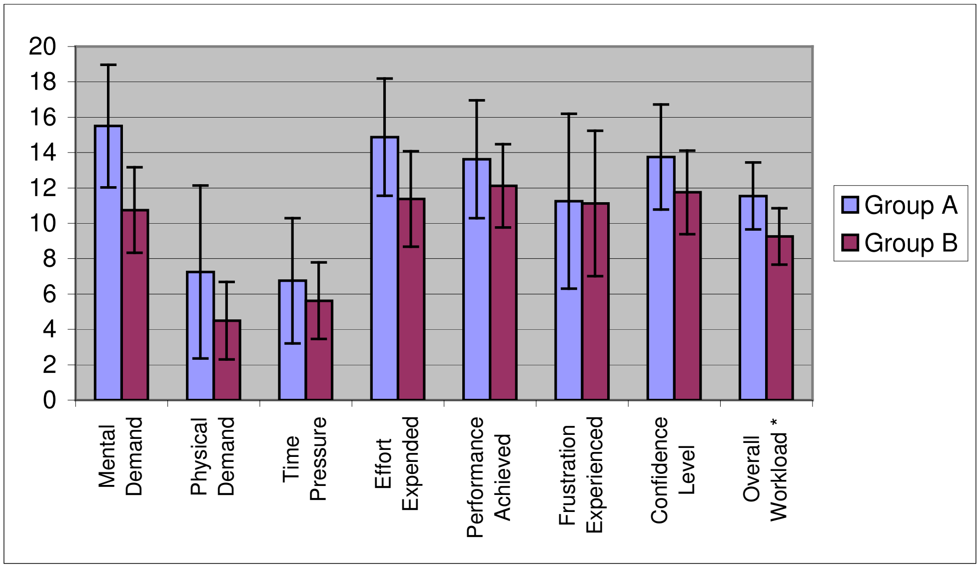


Figure 42. Results of workload analysis measured using a NASA TLX workload evaluation form showing mean scores. 'Performance Achieved' and 'Confidence Level' have been inverted such that a lower score indicates a lower workload. Error bars indicate a confidence interval of 95% about the mean.

Table 3 displays the results of a Mann-Whitney analysis of the workload factors. A significant difference was demonstrated in the 'mental demand' factor ($W_{14} = 88.0$, $p = 0.039$). Group A indicated a mean mental demand of 15.5, where as Group B indicated a mean of 10.8. The difference in perceived overall workload is also significant ($T_{14} = 87.5$, $p = 0.046$). The mean perceived overall workload for Group A is 10.33, compared to 8.54 for Group B. No other significant workload differences were noted in the individual factors.

Workload Factor	Group A Mean	Group B Mean	W	P
Mental Demand	**15.5**	**10.8**	**88.0**	**0.039**
Physical Demand	7.25	4.5	79.5	0.24
Time Pressure	6.75	5.63	72.5	0.67
Effort Expended	14.88	11.38	83.5	0.109
Performance Achieved	13.68	12.13	76.0	0.43
Frustration Experienced	11.25	11.13	66.5	0.92
Confidence	13.75	11.75	77.5	0.34
Overall Workload	**10.33**	**8.54**	**87.5**	**0.046**

Table 3. The mean scores for the individual and overall workloads. The results of a two-sample T-test are also included.

5.4.6 Discussion of the Virtual Versus Tradition Training Results

The results do not support Hypothesis 1. There was no significant difference found between any of the performance measures when comparing the two groups of participants. Although not significant, the mean number of follicles correctly position was slightly higher for Group A than for Group B. This is somewhat offset by the fact that Group A participants had a higher mean score for the number of extra follicles placed where none existed. Again although no significant difference was detected, Group A displayed a slightly higher mean score of other surface features mistaken for follicles. When sizing the follicles, Group A had means indicating slightly higher performance although the differences were shown not to be significant. There are a number of factors that influenced these results. Firstly, the simulator training session was not designed as a teaching aid, but more as practical experience. Participants were given no feedback about their performance, and therefore could not learn from their mistakes. Performance on the simulator would be expected to improve more rapidly if particular care was given to teaching methods such as presenting users with results from a training session either during or immediately after the session. Also, a weakness in the experiment was that the virtual and specimen ovaries were based on different models. Although the virtual models share many similarities with the specimen cow ovaries, their shape, size and haptic properties are more similar to horse ovaries. The range of follicle sizes that were used in the virtual ovary training sessions was between 2cm and 3.5cm in diameter. These sizes of follicles were similar to what would be expected on horse ovaries. In the bovine specimen ovary examination, follicles ranged from 0.5cm to 1.8cm in diameter. This could have affected the virtually trained group when estimating the size of follicles on the specimen ovaries. Finally, as virtually trained participants had not been exposed to specimen examination before, they had encountered virtual ovaries with follicles only. It should be expected that they would make more mistakes when distinguishing between follicles and other surface features.

Hypothesis 2 has again not been supported. Although there seems to be a large difference in mean times between Group A and Group B, there was no significant difference detected between the groups. The mean time taken for Group A was 24.9 seconds faster than the mean time for Group B for one examination that lasted no longer than 300 seconds. Two factors that contributed to this result not being significant were the high variance in the time taken for participants in both groups, and the examinations being capped at 300 seconds. Group B in particular displayed a wide range of mean times. Group B participants also had a mean of 2 examinations out of 8 stopped by the experimenter at the maximum time compared to 0.75 out of 8 examinations for Group A.

Hypothesis 3 was not supported by the results. In the case of 'Mental Demand', the results were the opposite of what was expected. Group A indicated a significantly higher mental demand than Group B. There is a general trend of Group A indicating higher workload factor scores than Group B, although none of the factors are significant. This is highlighted by Group A indicating a significantly higher level of overall workload than Group B. This again can be thought of as an indication that the group trained using the virtual methods subjectively found the task more difficult than the traditionally trained group. These workload results suggest that Hypothesis 3 was overly optimistic. Group A participants had only experience in the procedure through lectures and the simulator where as Group B had handled tracts before. Group B could therefore be expected to be more comfortable handling the specimen tracts than Group A.

5.5 Comparing the VR with the specimen results

Comparisons were also made between the performance on the virtual ovaries and performance on the specimen ovaries for the virtually trained group. As all virtual ovary examinations occurred before the specimen examinations, results from the final virtual training session are used as an estimation of virtual ovary examination skill at the time of the specimen examinations.

5.5.1 Hypotheses

5.5.1.1 Hypothesis 1

It is hypothesised that a correlation would be shown between a participant's performance in correctly positioning and sizing the follicles in a virtual examination, and a participant's performance in positioning and sizing the follicles in the specimen examination (to an accuracy of 0.5cm). The dependent variable for Hypothesis 1 is the performance in placing and sizing the follicles. The independent variable is the type of examination: either virtual or real.

5.5.1.2 Hypothesis 2

There will be a correlation for participants between time required to complete a virtual ovary examination and time taken to complete a real life examination. The dependent variable for Hypothesis 2 is the mean time taken to complete an ovary examination. The independent variable is the type of examination: either virtual or real.

5.5.1.3 Hypothesis 3

There will be a correlation in workload factors for participants between the different conditions. Particularly important to this analysis will be 'Mental Demand', 'Performance Achieved', and 'Confidence'. The dependent variables for Hypothesis 3 are the workload factors collected from the TLX workload analysis form. The independent variable is the type of examination: either virtual or real life.

5.5.2 Results

A Pearson product moment correlation test [21] was used to test Hypothesis 1. Figure 43 shows a graph of performance placing and sizing follicles in the virtual case against performance placing & sizing follicles in the specimen case. The results of seven participants were used to test this hypothesis as one participant's results were discarded due to reasons explained in section 3.5.1. The Pearson correlation coefficient is equal to 0.694 (p = 0.084).

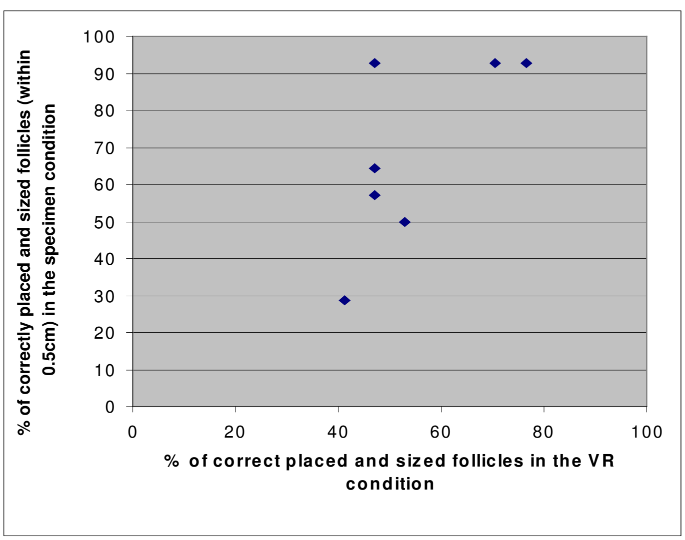


Figure 43. Graph of percentage of correctly placed and sized follicles in the virtual condition against correctly placed and sized with 0.5cm follicles in the specimen condition for all participants.

A Pearson product moment correlation test was also used to test Hypothesis 2. Figure 44 shows a graph of mean time taken to complete an examination in the virtual condition against mean time taken to complete an examination in the specimen condition. The results of all eight participants were used to test this hypothesis. The Pearson correlation for coefficient is equal to 0.04 (p = 0.93).

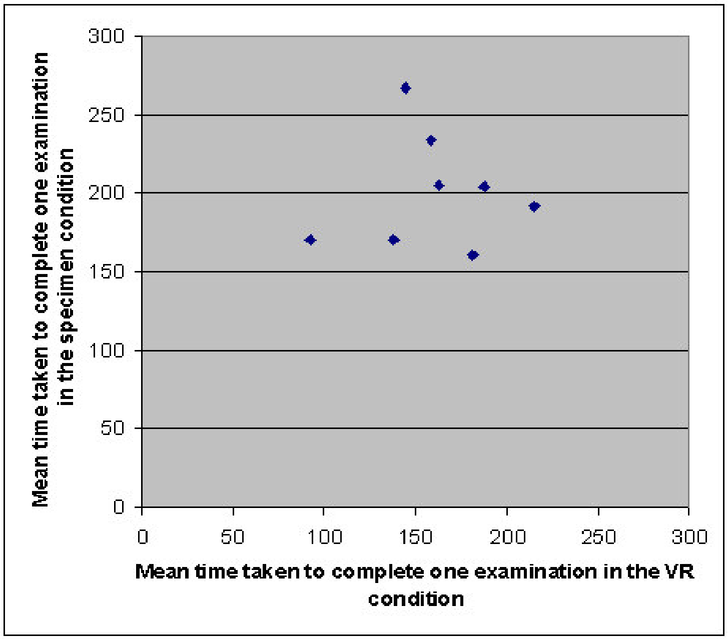


Figure 44. Mean time taken by each participant in both the virtual and specimen conditions. The maximum time taken for each examination was capped at 300 seconds.

Each of the different workload factors was analysed using the Pearson product moment correlation test. Results are shown in Table 8. The significant probability values are highlighted in bold type.

Element	VR Mean	Specimen Mean	Correlation Coefficient	p
Mental Demand	9.6	15.5	0.65	0.078
Physical Demand	**3.3**	**7.3**	**0.76**	**0.029**
Time Pressure	2.8	6.8	0.67	0.070
Effort Expended	10.6	14.9	0.53	0.174
Performance Achieved	8.5	13.6	0.23	0.580
Frustration Experienced	**6.1**	**11.3**	**0.74**	**0.036**
Confidence	8.8	13.8	0.057	0.89
Overall	6.8	11.5	0.677	0.065

Table 8. Results of a Pearson product moment correlation test on the different workload factors. A p-value below 0.05 indicates a correlation between the measured VR and specmen mean workload factor.

5.5.3 Discussion

Hypothesis 1 is not supported by the results. A Pearson correlation component of 0.694 indicates a level of correlation but this is not significant although it tends towards significance (p = 0.084). This suggests that a larger participant pool might provide significant results.

Hypothesis 2 is again not supported by the results. A Pearson correlation component of 0.04 (p = 0.93) indicates strongly that there was no correlation between time taken for a virtual examination and time taken for a specimen examination for each participant.

Analysis of the workload data provided mixed results. There was a significant correlation between conditions for participants when ranking physical demand (p = 0.029) and frustration experienced (p = 0.036) of both conditions. For the other workload factors, there was no significant correlation. The factors 'Mental Demand' (p = 0.078), 'Time Pressure' (p = 0.070), and 'Overall Workload' (p = 0.065) were not significant, but tended towards significance, where as all other factors showed strongly no correlation.

5.6 Conclusions

This chapter has presented one of the most in depth studies in the literature to validate a medical simulator, in particular a palpation simulator. Several studies have been described in Chapter 3 that test the validity of surgical and MIS simulators, but there have been few attempts to validation a palpation simulator. The construct and concurrent validity of HOPS has been tested. The concurrent validity study took advantage of the availability of a Virtual Reality simulator to present multiple shorter training sessions to students rather than one long session. This would not be feasible through traditional methods.

A comparison of ovary examination between a virtually trained group and a traditionally trained group of students has been presented. Their performance during specimen ovary examination was studied as an initial attempt to validate the HOPS simulator as a training tool. As part of this experiment, a study has also been described that examines the performance of a group of students on the HOPS virtual simulator over several training sessions. The results from this section of the experiment showed that improvements were made by participants using the simulator in both performance on the simulator and time taken to perform an examination. Workload results also showed that as participants had more practice on the simulator, they found the task of identifying the virtual follicles easier. It is also important to note that these improvements were all down to practice on the simulator, as they did not receive any feedback on their performance on the virtual ovaries until after the specimen condition had been finished.

A further experiment was designed to examine if the performance improvements when using the simulator carried over to the real life ovary examination procedure. The hypotheses in this section of the experiment were not supported. No significant differences were found in the performance of the virtually trained and the traditionally trained groups on the specimen. Time taken for the examinations was again found to be not significantly different between the groups. Measured workload factors suggested that the virtually trained groups found the task harder mentally. These results suggest the initial hypotheses were too optimistic.

This experiment however showed that both groups displayed similar performance on the specimen ovaries. Group A only had previous exposure to virtual ovaries where as Group B took part in a standard two hour long practical laboratory session on the reproductive system of a cow. It is encouraging that there are similarities in the performance of the virtually trained group with the traditionally trained group. This suggests that the training tool might be particularly beneficial for the course in situations where resources restrict the use of traditional training methods.

This chapter has not been able to provide a strong validation for the HOPS simulator in order to answer the second key research question positively. However some encouragement can be taken from the fact that students who had no previous exposure to specimen tracts performed similarly to students who performed the standard anatomy lab training session with similar tracts.

6 Augmenting Virtual Medical Training

6.1 Introduction

Presenting performance feedback to users is an essential feature of a simulator [44]. Without feedback on performance, novices will not be able to identify and therefore correct errors that they make. Chapter 4 has discussed the difficulties that can occur when trying to provide feedback to a user during a procedure. Particularly in the case of ovary palpation where the examination is internal, and the teacher often finds it hard to guide the student, as the student might not know where he or she is in the animal or what structures he or she is feeling. The third research question described in Chapter 1 presents the possibly of using a Virtual Reality simulator as a tool for assessing performance. This chapter attempts to formalise methods for achieving this.

6.2 Description of the Multimodal Cues

6.2.1 Background

Recent research has concentrated on providing post procedure performance analysis, but there has been little work on guiding a user during the simulation. Providing users with multimodal cues during a simulated examination has the potential to both guide them, and present them with performance feedback during the simulation. Recent studies have shown that this is a potentially useful area for providing training in gestures. Feygin *et al.* [36] conduct a study into the possibility of providing gesture training using either visual, haptic, or visual-haptic guidance. Further to this, there were 2 conditions in which participants recalled the gestures. The conditions were haptic-visual where the participants saw their cursor as they attempted to perform the gesture, and haptic where the participants attempted the gesture with no feedback of cursor position. Participants performed each of the training conditions twice - once with haptic-visual recall and once with visual recall - making 6 conditions in total. Thirty six participants took part in a within groups study, where the conditions were performed in a counterbalanced order.

Training consisted of three practice trials to demonstrate the three different conditions. The experimental task was to perceive a gesture using one of the conditions above, and reproduce that gesture. Both spatial and temporal aspects of the gesture were important. One gesture was used for the study, however, this was rotated in order to create alternate gestures of the same complexity. In all, 6 different gestures were created by either rotating the gesture or inverting the initial trajectory. The experimenters claim that participants stated that they were unaware of any similarities in the gestures. For each condition, one of the six gestures was played to each participant twice. The participant then tried to recreate the gesture. This sequence of training and recall was repeated fifteen times for each condition. Results showed significant improvement in recreating the gesture in all conditions between the first and the last gesture. The haptic only training mode performed significantly worse than the haptic-visual training mode, but not significantly worse than the visual training mode. This study demonstrates that both visual and haptic training methods can be used to train users in performing short gestures. It still must be demonstrated however that these results would extend to longer movements and real world tasks. From a real world training perspective, it would be important to show that the motions learned using one of these methods in the virtual world translated into improvement in performance in the real world task. This study does not take into account the possibility that contact with objects may be required to complete a task. This will have implications as to the use of haptic feedback as will be discussed in Section 6.2.2.

The following sections examine several possible multimodal cues that could be used to either provide training to a user during his or her examination, or allow a teacher to assess the performance of a user after the examination.

6.2.2 Haptic Cues

Haptic cues provide a method to directly affect the user's path through the environment. Different forms of guidance that haptic cues could provide have been considered. These are guidance through pre-recorded movements; and interactive guidance.

6.2.2.1 Guidance Through Pre-Recorded Movements

Guiding a user using pre-recorded movements can be broken down into 2 stages - record and playback. During the record stage, both positional and force information of a user must be stored at specific regular sample intervals. Playback of the procedure would drive the user's interaction point along the path recorded. At each stage, the driving force would equal the force recorded for the current position. By this method, a student could feel the techniques and forces involved in a correct procedure by playing back a recording of an experienced doctor or veterinarian performing the same procedure. This would be analogous to the student's hand being controlled by the veterinarians hand and dragged through the appropriate motions. In an ideal system, the student should still perceive the reaction forces from the contacted objects, however, Section 6.2.2.3 discusses why this may be difficult to achieve.

Dang *et al.* [24] discusses a training system that provides guidance to users by restricting their movements from deviating from a path. This method allows a user to follow the path taken for a procedure by an expert, but allows the user to apply the forces to perform the surgery.

Alternatively, haptic cues could be used by an expert to assess the performance of a student by playing back a recording of his or her movements. The expert could follow the path of the student to gain an overview of the student's examination.

6.2.2.2 Interactive Guidance

Interactive guidance can be thought of as a tutor-trainee model similar to the traditional teaching technique for medical training. In this situation two interaction points would exist in the same environment. The first is controlled by the trainee, and can be used to explore the environment freely as in a single user simulation. The second is controlled by the tutor, who can influence the student at any time. This could take the form of grabbing the student's interaction point and dragging it through a series of motions. The trainee could then practice the procedure as normal, while the tutor could guide him or her if and only if help is required. This would serve to reinforce the apprenticeship model currently in use, while allowing the tutor to have a more active role in guiding the student. This is analogous to a dual control car that is often used when teaching learners to drive. The instructor allows the student to control the movements of the car, only intervening when necessary, to correct an error from the student for example.

6.2.2.3 Implementation

Implementation of a haptic guidance system is not trivial. Using HOPS as an example, during the recording stage, the position of the PHANToM can be sampled at a rate of up to 1000 Hz. The PHANToM's position sensors provide a representation of the current cursor position that can be used to accurately recreate the path recorded. However, recording force information at the sample points presents problems, as the PHANToM device does not have force sensors. A different haptic device with force sensors to sense the user's force would avoid this problem. The system implemented attempts to estimate the applied force through the reaction force from objects or effects within the scene. Using this system however, no forces would be present to guide the user when the recording was moving through free space as the reaction force measure would be 0N. By introducing viscosity throughout the scene, a reaction force to any movement can be detected through the device.

Playback also presents problems, as even a passive user can affect the path of the cursor, and applying the recorded force vector will not generally move the cursor along the recorded path. The PHANToM drags the user's finger through a series of movements. Resistive forces from the user will combine with the driving force to produce deviations from the path. The naïve system would calculate the magnitude of the recorded force and apply it towards the next sample point on the path. This however can cause instabilities when contacting objects in the scene, demonstrated in Figure 45. Since the instabilities are due to the conflict between the playback force and the contacted object's reaction force, if the contact point is very close to the surface the object – and therefore the reaction force is low – the user may feel a slight buzz. However, severe problems occur when further inside an object. In this case, the user would feel and hear a strong vibration when contacting an object that would obscure the feel of the object and therefore negate the usefulness of the haptic cue. This situation constantly occurs when using forces similar to those used an ovary palpation procedure, even though the objects are modelled as soft objects.

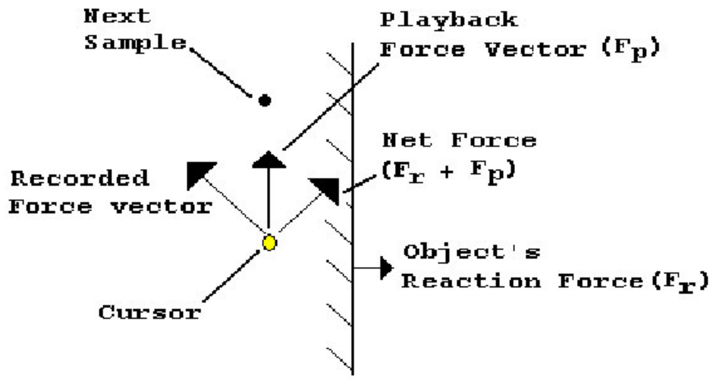


Figure 45. Demonstration of the problems of playback when the cursor and next sample point lie inside an object. The direction of the playback force will be affected by the object's reaction force and resultant force will push the cursor away from the sample point.

Figure 45 shows a situation in which the user's cursor is inside an object. The next sample point is also shown to be inside the object. The user feels the playback force from the haptic cue (Fp) pulling him or her towards the next sample point. However, the path of the user is also affected by the force felt due to the fact that he or she is in contact with an object (Fr). These force vectors combine such that the force the user feels will push him or her closer to the surface of the object than was intended. This force vector is shown as Net Force in Figure 45. At the next update, the direction of Fp will be adjusted to compensate for the deviation but again will not account for the new reaction force felt by the user.

During an examination, when touching an object the user will apply a force to counteract the reaction force from the object. Even if the user is moving perpendicular to the surface of the object, he or she is still applying a force to counter this reaction force. However, the playback force as described above does not take account of this and attempts to drive the user directly towards the next sample point. A more complex algorithm would adapt the playback force direction depending on the reaction force from the contacted object. This force would need to have an extra component that is equal in magnitude but opposite in direction to the object's reaction force. The result of this, however, is that the user would not be able to feel any force due to penetration depth of the cursor within the object and therefore would not be able to feel the virtual object.

6.2.3 Audio Cues

Audio cues can be used to convey state information about a system to the user. Gaver presents ARKola [39], a system in which continual auditory feedback can be used to monitor the running of a simulated factory, and in particular alert the user when an error occurs. Audio feedback to present state information has also been demonstrated in a medical context to augment a surgical procedure [120]. Surgical instrument position and optimal path information are passed to the surgeon through audio, allowing a surgeon to use the information while keeping his or her visual focus on the patient. Similar concepts of supplying users with auditory position and path information can also apply to medical simulators. Incorporating audio warnings into a simulation can provide immediate feedback to users that the current action they are performing is incorrect, or dangerous. Unlike haptic cues, audio cues do not directly affect the cursor position, but allow users to correct their actions themselves. In this manner, they can build confidence as they progress through the procedure that their actions are not damaging the patient.

Alternatively, audio cues could be used to train novices in the forces to be used for a procedure. The pitch of an audio tone could be varied depending on the force used during an examination. A novice could listen to the change in pitch as an expert contacts an object, and can then attempt to match this audio profile during his or her examination.

6.2.3.1 Implementation

A simple audio warning cue has also been incorporated into HOPS. One of the dangers when performing an ovary examination is damage can be caused by palpating an ovary too firmly. An obvious solution is to alert the user when he or she is applying too much pressure to an ovary. Each of the objects in the environment can be assigned a threshold force value. If more than this threshold force is applied to an ovary during palpation an audio warning is sounded. This warning sound could be abstract to the situation (like a constant tone), or might be designed to fit in with the context such as using the appropriate animal noises for a veterinary medical simulator. In the current implementation, the pitch of an abstract audio warning is linked to the force applied above the threshold, so a higher pitch of sound indicates a greater danger.

6.2.4 Graphical Cues

Graphical cues can be most easily used to highlight a specific area of the screen. The user's attention can be drawn to a particular object by colouring it differently from the surroundings. In this way, an area of interest, for example an incision point, can be highlighted.

Other possibilities exist however. A training system might include a path following mode where, in much the same way as described in Section 6.2.2 for pre-recorded haptic cues, a pre-recorded procedure is played back to the user. During a correct examination, cursor position could be recorded at regular time intervals and a representation of this position can be played back in a subsequent examination for a student to follow. The pre-recorded cursor would provide no direct guidance, but a user could follow the movements with his or her own cursor, performing the same actions as the recording.

Alternatively, graphical cues can be used to assess a student's performance on the simulator. The following sections describe the implementation and evaluation of a system to assess the performance of a simulated examination using HOPS.

6.2.5 Natural Cues

The above cues are designed to provide extra guidance over and above what could be provided in the real life procedure. Natural cues can be defined as cues that the user would be expected to receive during the real life procedure that would provide information about some feature of the procedure. One example in the literature of such cues would be the Preop endoscopic simulator [14] where the virtual patient will cough during the bronchial examination if not anaesthetised properly. These cues are important for a student too learn so that they can be recognised in the real life procedure, and the student can learn to adapt his or her behaviour based on these cues.

For a virtual ovary palpation procedure the horse could move when uncomfortable. For example, if the examiner is using too much force during his or her examination, or has spent too much time examining the horse. A user of the simulation would feel this haptically through the PHANToM as his or her frame of reference changed.

Auditory feedback can be used in a similar manner to provide cues the user. Again the animal may make noises during an examination that can influence the behaviour of the veterinarian. For example, if the animal is uncomfortable, the appropriate sound could be played to the user.

Olfactory cues are an important factor in some examinations that are often overlooked. This is almost certainly due to the immaturity of the technology and the difficulty in generating a range of different smells convincingly.

6.2.6 Combining Cues

One possibility not yet discussed is the potential for using combinations of cues when guiding the user through the procedure. This could be used to provide complementary information to the user through the different modalities. For example, graphical cues provide a good method for allowing the user to follow through the different techniques involved in completing a procedure. They may not, however, provide a reliable method of providing the user with a sense of the contact forces required during an examination. This could be remedied by combing a graphical playback cue with an auditory cue where pitch is linked to the contact force. The user could therefore follow the cursor visually while trying to match his or her current auditory pitch with a pre-recorded sound from an expert.

Alternatively, using combinations of cues could be used to avoid some of the problems associated with a particular cue. For example, haptic cues provide a good method of guiding a novice through an examination. However, haptic cues during contact with an object have already been highlighted as problematic. In a combined cues system, haptic guidance cues could be used to drag the user through free space. When contact occurs, the cueing mechanism could switch to a combination of graphical and auditory cues allowing the user to follow the cursor without the instabilities associated with the haptic playback cues.

One obvious situation in which combining different modalities of cues would be beneficial is in a natural cueing environment where an action might have a consequence that leads to a series of different modalities of cue being required. One obvious example is in situations where the patient may experience pain or discomfort. This may lead to involuntary movements such as those displayed by the Preop simulator (described in 6.2.5). The user would feel his or her frame of reference change as the virtual patient coughed. An auditory cue would also be heard in the real life procedure. The combination of cues in this instance would provide the user with a convincing representation of the situation in the real life procedure. Without the sound in this instance, the visual feedback given to the user during a cough may become unstable for short periods without explaining to the user why this was happening. It could possibly even be dismissed as a glitch in the system.

6.2.7 Overview

The multimodal cues described allow the user's of a Virtual Reality medical training simulator to provide training and assessment in a manner that is not possible in traditional training environments. These have the potential to be a useful addition to Virtual Reality training systems. However, it must be demonstrated that they do provide some benefit to the user in either improving his or her performance, or allowing assessment of his or her current skill level. The study by Feygin *et al.* [36] described in Section 6.2.1 shows encouraging results. However, further work is required in a more realistic context before their usefulness can be assessed. For this reason, a study will now be presented that evaluates the use of graphical feedback cues to assess the performance of a user on a Virtual Reality medical simulation.

6.3 Case Study: Graphical Playback Cues for Assessing Ovary Examination Performance

6.3.1 Background

Section 3.5 discusses a growing need to be able to assess the performance of a medical clinician. Michell [77] describes how assessing clinical competence is an important area in both human and veterinary medicine, and argues that regular competence testing and certification of doctors for specialist procedures will soon be an important part of medical training. A system for assessing student performance would also be beneficial for veterinary schools. From Chapter 4, it is clear that there are difficulties in teaching and assessing a student's performance during an equine ovary examination.

This study examines the use of graphical cursor playback to assess performance in a medical virtual environment. Although this is not a novel technique, there have been no studies to date to judge its effectiveness as an assessment tool in the medical domain. Burdea *et al.* [16] present an example of a prostate palpation simulator that included a graphical playback feature although there was no discussion of its usefulness. The playback consisted of a simple cursor that moved through various recorded sample points and allowed playback at the same rate as the initial examination. Feygin *et al.* [36] present a study that examines the use of different forms of playback for training users to perform particular gestures in three dimensions. This is discussed in more detail in Section 6.2.1. Feygin's method to analyse the results was determined through a closeness of fit to the original gesture. This can be seen as a form of assessment of a gesture using a recorded cursor trace to provide details after the gesture has been performed.

This method is also commonly used when assessing the technique of a sports person. Limb positions can be monitored at regular intervals during a sports action using body tracking equipment. Commercial systems such as the MotionMonitor developed by Innovative Sports Training [50] allow recording and visualisation of a sports action for training purposes. Motion capture equipment is used to develop a three dimensional computer model, which can then be compared to an ideal technique. This study examines an expert's ability to assess recorded ovary palpation examinations using the Glasgow Horse Ovary Palpation Simulator.

The importance of performance feedback to students has been previously established in Chapter 3. This study has chosen to look initially at feedback given by experienced veterinarians who are involved in teaching veterinary students to palpate horse and cow ovaries. If successful, future studies would examine the possibility of providing the feedback directly through the computer by examining the recorded examination for characteristic mistakes. This will be discussed further in Section 6.8.

6.3.2 Developing the Experimental Environment

6.3.2.1 Developing the Recorded Examinations

The initial stage of the experiment was to develop a range of examinations of differing quality. This was done in close collaboration with a veterinarian who teaches at Glasgow University Veterinary School, and is experienced in ovary palpation. Common mistakes made by novices were identified by this veterinarian. The identified factors that affected the quality of an examination are described below:

1. *Search Strategies* - Whether the examiner used a logical search pattern of the ovaries to look for surface features, or whether he or she randomly searches the ovaries. This was measured on an equal appearing interval scale of ten values with 'Random' and 'Ordered' marked at the extremes.

2. *Missed sections of the right ovary* - The examiner may not have examined the whole surface of the ovary. There may have been sections of the ovary that were not palpated during the examination that may have contained surface features. This was measured on an equal appearing interval scale of ten values with 'None' and 'Thorough' marked at the extremes.

3. *Missed sections of the left ovary* - This is as described in 2, but for the left ovary. This was again measured on an equal appearing interval scale of ten values with 'None' and 'Thorough' marked at the extremes.

4. *Follicles palpated* - The examiner may have searched the ovary thoroughly but failed to find a surface feature. He or she may have touched a follicle, but did not identify it, and therefore did not try to size it. This was measured as yes or no answers for each follicle depending on whether they have found and identified the follicle or not.

5. *Pushing too hard* - The examiner may have used excessive force during the examination that would cause damage or distress to the animal. This was measured as the number of occasions that excessive force had been used and grouped into categories. Either no times, 1 to 3 times, 4 to 6 times, or above 6 times.

Five examinations performed by this expert veterinarian were recorded. Of the five examinations, deliberate errors were introduced into four of them to test the different features of the system. Categories 2 and 3 both referred to missed sections of an ovary. It was therefore decided that inclusion of only one examination that tested either category 2 or 3 was necessary. The vet who performed the recorded examinations was asked to introduce one of these features into each of four of the examinations. He was not instructed on how to introduce the feature, but told for example to provide an examination where the follicle on the front bottom of the left ovary was not identified during the examination. The examinations developed were designed as follows:

- E_{gd} - An examination was developed where an ordered, thorough search of the ovaries was performed. All follicles were found and identified, and the examiner rarely used too much force. This represented a good examination.

- E_{ss} - An examination was developed where a poor search strategy was used. The examiner moved between ovaries several times during the examination, and did not cover the ovary surfaces in an ordered manner.

- E_{tmf} - An examination was developed where too much force was used consistently. The examiner pushed above the defined safety threshold on several occasions, and on different objects within the scene.
- E_{mo} - An examination was developed where the examiner failed to identify the right ovary. A thorough search of the left ovary was performed. Although the right ovary was contacted on some occasions, the examiner did not identify that this was the case as these contacts were very brief.
- E_{mf} - An examination was developed where two of the three follicles on the ovary surface were not identified. Although the examiner came into contact with all the follicles, only one was identified and therefore only one was sized.

Screen shots of the full path of each of the examinations can be seen in Appendix E.3.

6.3.3 Graphical Feedback Cues

During an examination using HOPS, it is possible to record information about the actions of the user. In this case, the information stored is the position of the cursor, and the reaction force from the object currently touched.

The graphical feedback cues take the form of spheres drawn at points sampled several times a second. During the playback, the drawing mode for the ovaries and follicles is changed to display only a wire-frame representation of the structures. This is to allow participants to view the cursor as it pushes into or behind a structure. The colour of the follicles has been altered from red to pink to enable a user to distinguish between the two more easily.

The colour of the current playback sphere is linked to the reaction force from the currently touched object. An example of an examination being played back is shown in Figure 46.

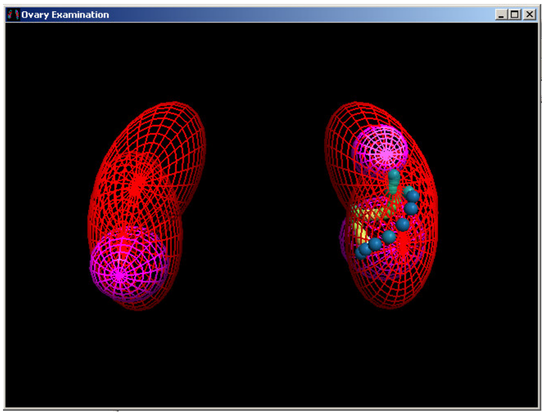


Figure 46. A pre-recorded examination that is being played back on HOPS. The purple sphere indicates the current position of and reaction force applied to the cursor. The spheres indicate the recent path of the examiner with the colour of the spheres varying as the force used changes.

A 'ghost' trace of the cursor is also available. This shows the recent cursor behaviour by displaying the previous 20 sample points as shown in Figure 46. This leads to the effect of a cursor with a 'tail' of sample points. Again, the colour of the sample points forming the tail is altered according to force applied.

The colour scale used is shown in Figure 47 and is a subset of a rainbow colour scale described by Levkowitz [68]. This scale has been developed to be perceptually linear, which means that a colour twice as far along the scale will be perceived to be representing twice as large a value. In the experiment, a blue colour cursor shown at the left of the scale represents a recorded sample point where no reaction force is being felt by the examiner. As the cursor colour moves to the right on the scale, the reaction force value from the object at that cursor position during the examination is higher. If the reaction force exceeds the predetermined safety threshold of 1.44N, the cursor colour becomes white as shown to the right of the colour scale in Figure 47. The scale used is made up of 130 distinct colours. Increments in colour therefore occur every increase of 0.011N increment in force. The actual value of the safety threshold was not important for the purposes of this experiment. The value of 1.44N was chosen as estimate of an appropriate value.

Figure 47. The rainbow colour scale used to indicate changes in reaction force during the examination. As the reaction force increased, the colour of the cursor moved linearly along the scale from left to right. If the reaction force exceeded the predetermined safety threshold at the very right of the scale, the cursor colour changed to white.

6.3.4 Controlling the Playback

Users interact with the system using the dialog box shown in Figure 48. For the purposes of this study, the participants did not use the load feature as loading of files was performed by the experimenter. Users did not interact with the timer section of the dialog box as this is only used to display timing information to the user. The display shows the length of the examination in seconds and the current position in the examination.

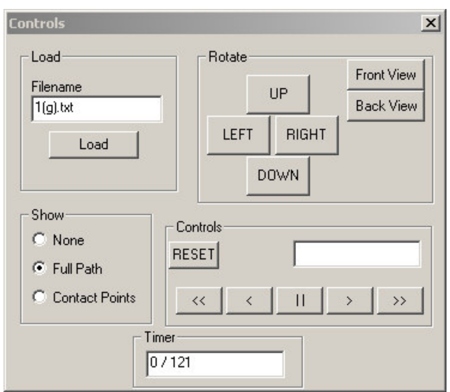


Figure 48. The control dialog box for the experiment. Users could rotate the scene, play or rewind at various speeds, or pause the examination. They could also choose to view the full examination all at once, or just the contact points.

A user can control his or her view of the playback using the controls in the 'Rotate' section of the dialog box. The buttons labelled 'LEFT', 'RIGHT', 'UP', and 'DOWN' rotate the scene in the appropriate direction around the centre point of the scene. The buttons labelled 'Front View' and 'Back View' provide a method of directly moving between the default start position, and directly behind the ovaries. Rotation could also be achieved by pressing and holding the left mouse button on the viewing window and moving the mouse. Horizontal mouse movements resulted in rotation of the scene about the Y-axis corresponding with the 'LEFT' and 'RIGHT' buttons on the control dialog box. Vertical mouse movements resulted in rotation of the scene about the X-axis corresponding with the 'UP' and 'DOWN' buttons.

The playback of the examination could be adjusted in the 'Controls' section of the dialog box. The buttons shown operate similarly to a video recorder in that a user can play, fast forward, play backwards, rewind, and pause the examination. The 'RESET' button returned the examination to the start. The 'Fast Forward' and 'Rewind' options allowed users to view the examination at twice the speed either forwards or backwards.

The 'Show' section of the dialog box controls what sections of the examination are viewed at any one time. There is the option to view all sample points in the examination at the one time. The result of viewing the whole path for one examination is shown in Figure 49. In this situation, a sphere is displayed on the screen for every sample point in the examination. These sample points have colourings that are linked to the reaction force in the same manner as described above.

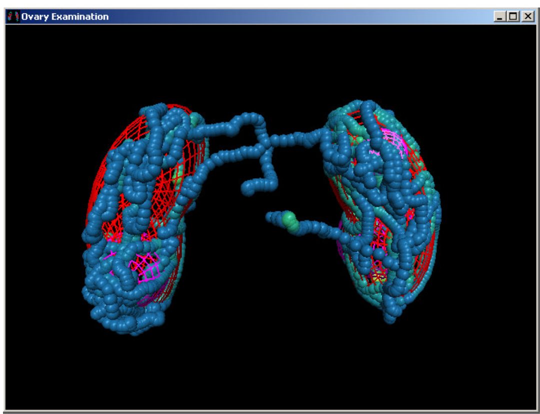


Figure 49. A pre-recorded examination that is displayed in its entirety on HOPS using the show full path option. Colour of the spheres indicate the reaction force on the cursor at that position.

Alternatively, a user can choose to view only the sample points where contact is made with one of the structures in the scene. An example of this is shown in Figure 50. This will remove the information where the examiner is searching, or moving between ovaries. However, if the assessor is currently interested only in the areas of contact, it can remove some unnecessary clutter from the screen.

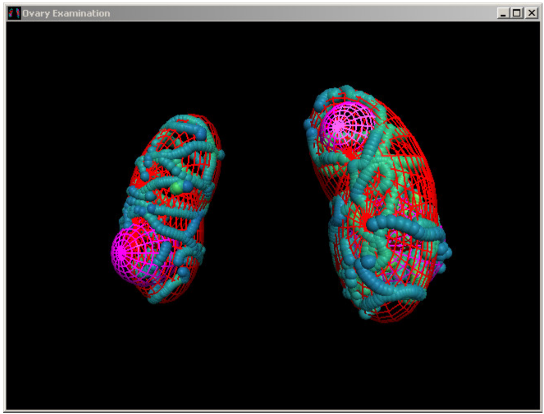


Figure 50. A pre-recorded examination that is displayed by HOPS using the option to show contact points. All points during the examination where contact with one of the structures occurs are shown.

The options to view the full path, or the contact points present only the spatial and force information, while removing the temporal aspects from the examination. They are intended to allow the examiner to get a quick overall view of the entire examination.

6.4 Method

6.4.1 Participants

One group of participants took part in the experiment. This group consisted of 10 veterinarians from a veterinary surgery. Practitioners from the surgery are involved with training students from Glasgow University Veterinary School in equine and farm animal procedures.

6.4.2 Assessing the examinations

Participants were presented with a visual only representation of the HOPS environment, shown in Figure 51. Each participant viewed the examination while controlling the graphical feedback that they received from the simulation. The same environment was used for all examinations. There were follicles on the front of the bottom half of the left ovary, and on the front of the top, and back of the bottom of the right ovary.

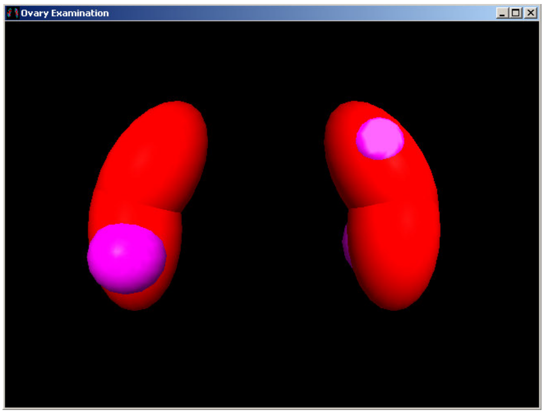


Figure 51. The Horse Ovary Palpation Simulator. This environment consists of a left and right ovary. Two follicles are shown as pink spheres in the front of the left and right ovaries. A further follicle exists on the back of the right ovary.

Each participant was presented with the examinations in a random order. He or she was asked to state his or her view on the quality of different factors of the examination. The factors assessed have been described in Section 6.3.2. The participant was also asked to give an overall rating for the examination on an equal appearing interval scale of ten values from 'Poor' to 'Excellent'. Participants were asked to rate their confidence that their answers are correct on a ten value unmarked scale from 'Not Confident' to 'Confident'. Workload data was collected after the task using the NASA TLX workload evaluation technique with added 'Fatigue' scale as described in previous experiments. A 'Confidence' scale was not included in the post experiment workload analysis as it was included in the answer sheet for each experimental case.

6.5 Hypotheses

6.5.1.1 Hypothesis 1

It is hypothesised that E_{gd} will be given a significantly higher overall rating than all other examinations. The dependent variable is the 'Overall Rating' for an examination. The independent variable is the examination being assessed.

6.5.1.2 Hypothesis 2

The dependent variable is the 'Search Strategy' used for an examination. The independent variable is the examination being assessed. It is hypothesised that E_{ss} will have a significantly poorer search strategy rating than all other examinations.

6.5.1.3 Hypothesis 3

It is hypothesised that E_{mo} will have a significantly poorer right ovary exploration rating than all other examinations. The dependent variable is the 'Right Ovary Exploration' coverage in an examination. The independent variable is the examination being assessed.

6.5.1.4 Hypothesis 4

It is hypothesised that participants will be able to distinguish whether a follicle has been identified or not with a high degree of accuracy.

6.5.1.5 Hypothesis 5

It is hypothesised that participants will be able to distinguish how many times the force safety threshold has been exceeded with a high degree of accuracy.

6.6 Results

6.6.1 Performance Results

Initially, the data were tested to check that they were normally distributed. The Anderson-Darling test was used in this instance [21]. The test was performed on the data for 'Overall Rating', 'Search Strategy', 'Right Ovary Exploration' and 'Left Ovary Exploration' scales. In each case, the data were shown not to be normally distributed ($p < 0.05$). It was therefore not possible to use ANOVA's to analyse the data from any of these scales. The Kruskal-Wallis test was used instead to analyse data from all four scales [21]. When a significant effect was detected by the Kruskal-Wallis test, a paired Wilcoxon test [21] was used to perform *post-hoc* comparisons on each of the individual results to discover where the effect occurred. This is the non-parametric equivalent of a paired T-test and allows significant differences to be detected in paired data sets.

Using the Kruskal-Wallis test, examination assessed was shown to have a significant effect on 'Overall Rating' of the examination ($p = 0.001$). *Post-hoc* comparisons were carried out using Paired Wilcoxon tests. Figure 52 shows the mean value given by the assessors for the 'Overall Rating' of each examination. Error bars from the graph indicate the confidence intervals of the data.

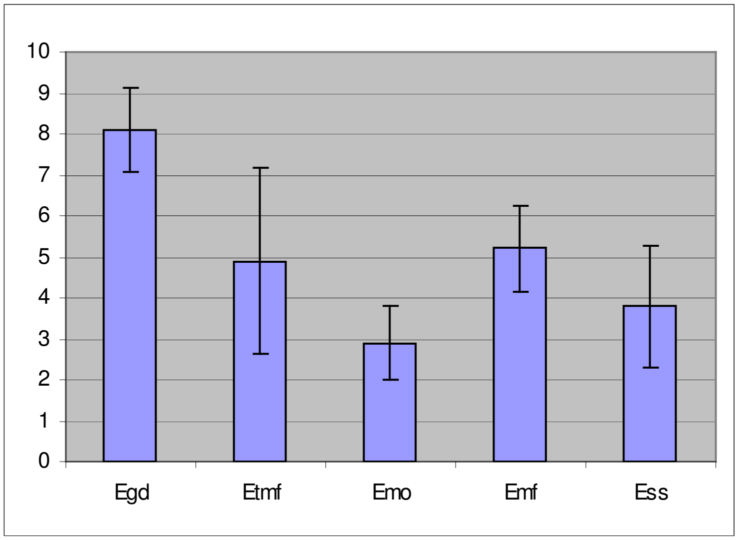


Figure 52. This graph shows the mean 'Overall Rating' given for each examination. The error bars indicate the confidence interval about the mean for each examination for a probability value of 0.05.

Significant differences were detected between the overall rating of examinations E_{gd} and E_{ss}, ($T = 53.5$, $p < 0.01$) E_{gd} and E_{mo} ($T = 55.0$, $p < 0.01$) and E_{gd} and E_{mf} ($T = 45.0$, $p < 0.01$). The difference in overall rating tended towards significance between E_{mo} and E_{mf} ($T = 8.0$, $p = 0.053$).

Similarly, a Kruskal-Wallis test was performed on the data gathered from the 'Search Strategy' scale. Examination was shown to have a significant effect on the rating given for a search strategy used ($p < 0.001$). *Post hoc* comparisons were therefore carried out using paired Wilcoxon tests. Figure 53 shows the mean values returned for each of the examinations for the 'Search Strategy' scale. Error bars indicate the confidence interval for each value. A significant effect was noted between E_{gd} and E_{mo} ($T = 55.0$, $p < 0.01$), and E_{gd} and E_{ss} ($T = 55.0$, $p < 0.01$). A significant effect was also detected between E_{mo} and E_{mf} ($T = 55.0$, $P < 0.01$).

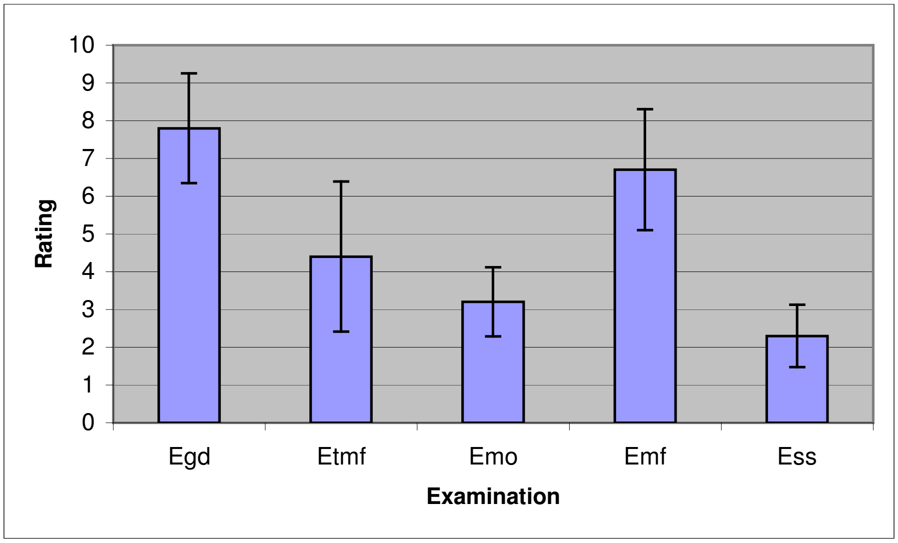


Figure 53. This graph shows the mean 'Search Strategy' rating given for each examination. The error bars indicate the confidence interval about the mean for each examination for a probability value of 0.05.

A Kruskal-Wallis comparison was performed on the data gathered from the 'Right Ovary Exploration' scale. There was shown to be a significant effect of examination on right ovary exploration ($p < 0.01$). Paired Wilcoxon *post hoc* comparisons were performed to discover where the effect occurred. Figure 54 shows the mean values returned for each of the examinations for the 'Right Ovary Exploration' scale. Error bars indicate the confidence interval for each value. A significant effect was detected between E_{gd} and E_{tmf} ($T = 40.0$, $p < 0.05$), E_{gd} and E_{mo} ($T = 55.0$, $p < 0.01$), an E_{gd} and E_{ss} ($T = 44.0$, $p = 0.013$). Significant differences were also found between E_{tmf} and E_{mo} ($T = 43.5$, $p = 0.015$), E_{mo} and E_{mf} ($T = 55.0$, $p < 0.01$) and E_{mo} and E_{ss} ($T = 2.5$, $p = 0.021$).

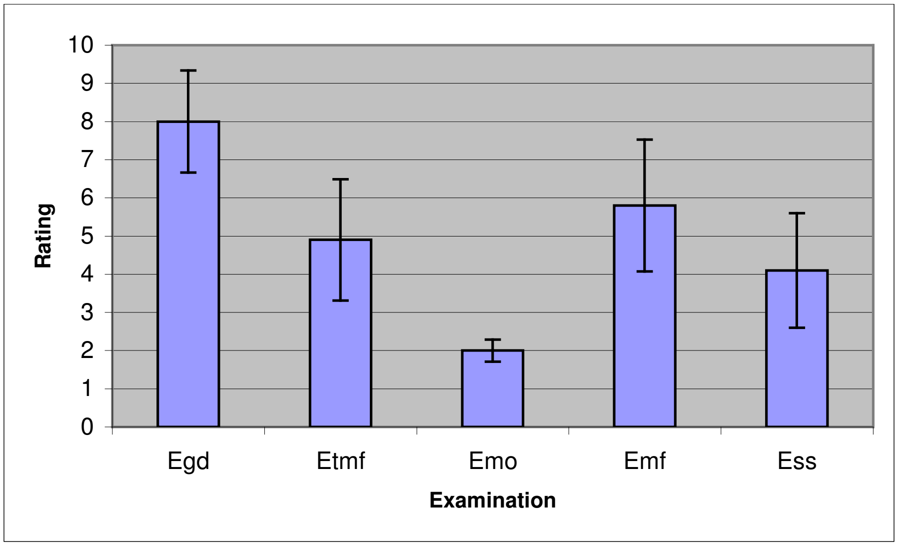


Figure 54. This graph shows the mean 'Right Ovary Exploration' rating given for each examination. The error bars indicate the confidence interval about the mean for each examination for a probability value of 0.05.

For completeness, a Kruskal-Wallis test was performed on data gathered from the 'Left Ovary Exploration' scale. The examination was found not to have a significant effect on left ovary examination ($p = 0.287$).

There were 15 follicles in total in the five examinations, with 9 follicles identified and 6 follicles not identified by the examiner. As all participants assessed all five examinations, 150 follicles were therefore present in total. Of these follicles, 90 were successfully identified and 60 were not identified by the examiner. Of the identified follicles, participants correctly answered that the follicle had been identified for 69 out of 90 follicles. A correct answer was therefore given for an identified follicle in 76.7% of these cases. For follicles not identified by the examiner, participants correctly responded for 49 out of 60 follicles. A correct answer was therefore given for a follicle not being identified in 81.7% of these cases. For all follicles examined, a total of 118 out of 150 were correctly answered by the participants. This means that participants were correct on 78.7% occasions.

For each examination, participants were asked to classify how many times the force threshold had been exceeded by choosing one of four categories. For the 50 cases examined by the participants in total, the correct category was chosen on 35 examinations. This means that 70% of responses were in the correct category.

6.6.2 Workload and Confidence Results

Participants were asked to rate their confidence that their answers were correct on a scale of one to ten. The results were analysed using a Kruskal-Wallis test. There was no significant effect of examination on confidence ($p = 0.559$). Overall, the mean confidence score for all participants for all examinations was 7.1 out of 10. Mean confidence for each examination is shown in Table 9.

Examination	E_{gd}	E_{tmf}	E_{mo}	E_{mf}	E_{ss}
Mean	7.7	6.2	7.6	7.1	6.8

Table 9. Mean confidence for all participants for each examination.

Workload results gathered from the participants after the task are shown in Figure 55. As this study involved only one group, no inferential statistics were carried out on these data.

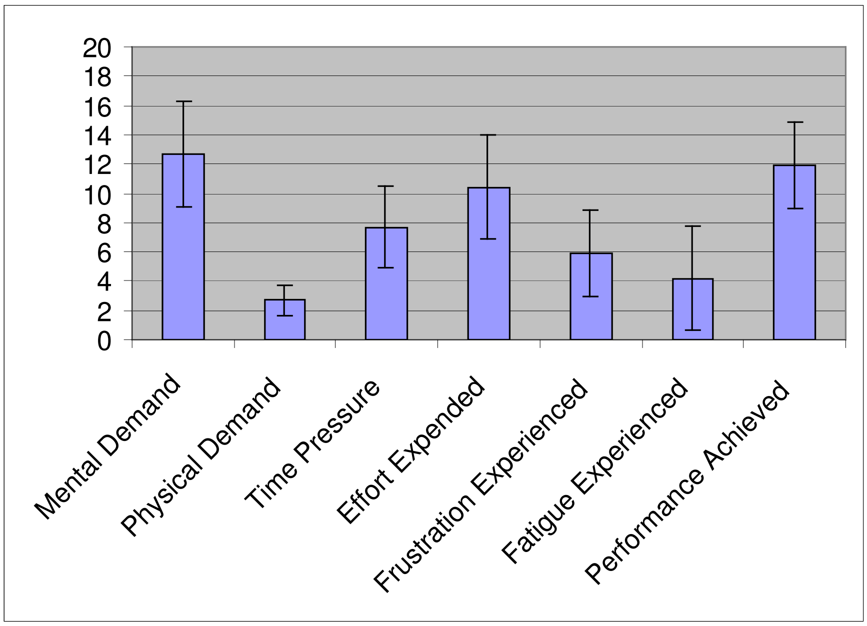


Figure 55. Mean score for all participants for each workload factor. For all scales except 'Performance Achieved', a lower score indicates less workload. For 'Performance Achieved', a higher score indicates a higher perceived performance

6.7 Discussion

Hypothesis 1 was to some extent supported by the results. E_{gd} was given an overall rating by the participants that was significantly higher than the ratings given to E_{mo}, E_{mf}, and E_{ss}. This means that participants judged E_{gd} as a better examination than E_{mo}, E_{mf}, and E_{ss}. Since E_{gd} was designed to be a thorough, ordered examination, and E_{mo}, E_{mf}, and E_{ss} were designed with mistakes, this shows some success for rating examinations using the visualisation technique. However, although the mean value of E_{gd} is higher than the mean value of E_{tmf}, the difference is not significant. The value of $p = 0.067$ does indicate that the difference is close to significance, but this could also be interpreted as due to the number of Wilcoxon tests being performed. The fact no significant difference was detected between these examinations can be explained by discussions with participants after the study had taken place. E_{tmf} was designed as an examination where the examiner uses too much force to palpate the structures. The majority of occasions where too much force was used were in contact with the ovaries or the bounding box representing the rectal wall. This is because the follicles are relatively small compared to the ovaries and the walls. Also, follicles are modelled as softer than both these structures the examiner needs to penetrate the structure further to achieve the same reaction force. Two participants indicated that they would not consider excessive force on the ovaries or rectal wall to be a problem, and it is excessive force on the follicles that is the biggest risk of damage, particularly in cows. Their overall rating for E_{tmf} was therefore not affected by excessive force on structures other than follicles.

Hypothesis 2 was again supported by the results to some extent. E_{ss} - designed as an examination with a poor search strategy - was recognised as having a poorer search strategy than E_{gd} and E_{mf}. A relatively ordered search strategy was used in both of these examinations although follicles were not identified in E_{mf}. However, there was no significant difference noted between the rating given to the search strategies for E_{ss} and E_{tmf}, and E_{ss} and E_{mo}. In the latter case, the result can be explained by the fact that the examiner failing to identify an ovary seriously affected the participants' perception of the search strategy for that case. This is understandable, as search strategy not only encompasses an examiners movement in finding the ovaries and moving between ovaries, but also his/her movements on the surface of an ovary. As there was little contact with the right ovary in examination E_{mo}, the assessor cannot say that an ordered search strategy has been used.

Hypothesis 3 was supported by the results. E_{mo} was an examination where the examiner failed to identify the right ovary, and was rated as having significantly worse right ovary exploration than all other examinations. Although this result might have seemed obvious from the descriptions of the examinations above, it could have proved challenging for the participants to determine when the ovary has been touched, and when the cursor is just close to the surface of the ovary. The difference only becomes clear when using the option to view contact points only. Figure 56 shows a comparison of a front view of the right ovary in full path mode and contact points mode for E_{mo}. Not all participants chose to use contact points mode, but could still make this distinction.

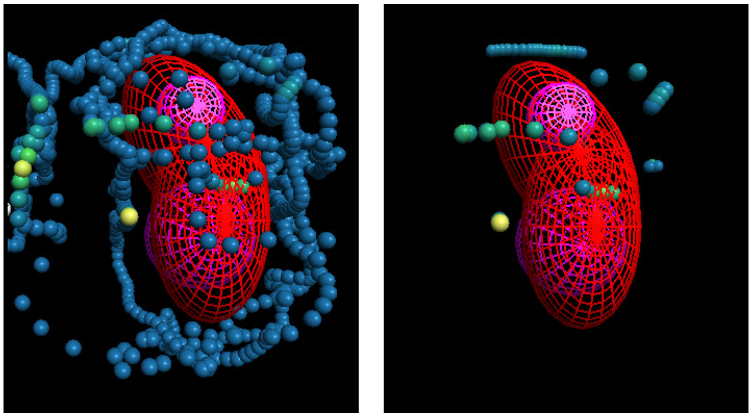


Figure 56. Shows a comparison of right ovary exploration between full path mode (left) and contact points only mode (right) for E_{mo}. The points shown on the right indicate either brief contact with the ovary or rectal wall.

Right ovary coverage was also judge significantly better in E_{gd} than in E_{tmf} and E_{ss}. One explanation for this could be that E_{gd} was judged to have a better search strategy than E_{tmf} and E_{ss}. When viewing the examination in play mode rather than all at once, using either full path or contact points mode, it may be easier to judge the thoroughness of an examination if an ordered search pattern is used. If the participant moves randomly over the ovary surface, it is more difficult to remember where he or she has been, and where he or she has still to explore. However, this is not borne out by examination of the data gathered through the 'Left Ovary Exploration' scale, where examination was determined to have no significant effect on the thoroughness of left ovary examination.

Participants were able to tell whether a follicle had been identified or not with a high degree of success. When a follicle had not been identified, participants could tell with an accuracy of greater than 80%, with over 75% of cases where the follicle was successfully identified being answered correctly. Overall, a 78% success rate was achieved with very little training. This would be expected to rise as participants become more familiar with the system.

Participants were also able to categorise use of excessive force into one of the four ranges with 70% accuracy. However, some participants mentioned the difficulty in distinguishing between the yellow used for high force, and the white used for excessive force. Despite this, participants were able to discern excessive force to good accuracy with very little training.

Analysis of confidence data for each examination showed that examination could not be shown to have a significant effect on confidence. Similar means for each examination are shown in Table 9, demonstrating that participants were similarly confident for each examination. As different examinations were designed with different faults, this might suggest that participants were equally confident when identifying different faults in an examination.

Although no inferential statistics were performed on the workload data, they can still be used to gain an insight into how challenging the participants found the task. The mean 'Mental Demand' measured for the task was 12.7, which indicated that participants did find the task challenging. Again this is emphasised by a mean value of 10.4 for the 'Effort Expended'. This is to be expected as the system presented to them is different from any system they have used before. Some of the participants did not regularly use computers. This would need to be taken into consideration in future systems, although the results shown above indicate that the task was successfully completed by the participants. Participants were under no pressure to finish the task within a specific time, so a low mean was expected for data gathered from the 'Time Pressure' scale. The mean value of 7.7 was surprisingly high. From observational results, some participants did feel under time pressure to complete each scale as the examination was playing in real-time. This restricted the length of time given for decision making in setting a value for each scale and led to the higher than expected value. This would become more important if this system was adopted into a veterinary course, where tutors may be assessing tens of examinations rather than just five. The mean perceived performance achieved was 11.9 out of 20. The results presented above would indicate that participants performed the task more successfully than their perceived performance. This could be explained by the participants finding the task challenging.

6.8 Future Developments

Participants indicated that in some instances it was difficult to decide on the part of structure being palpated, particularly when distinguishing the depth of the cursor position. This is understandable as a three dimensional environment is being represented on a two dimensional screen. To some extent, participants found that they could aid perception by rotating the scene. One solution would be to display surface contact points rather than actual cursor position when in contact with an object. In the current system, information is duplicated as reaction force - and therefore cursor colour - is dependent on penetration distance of the cursor inside the object. Similar information could therefore be presented by displaying surface contact points instead of actual cursor position inside the object. Users could then identify cursor position from the interaction between the cursor and the object's surface. If the lines representing the surface are drawn over the cursor then the cursor must be behind the object.

Although results suggest participants could successfully identify occasions where excessive force was used, some participants indicated difficulty in distinguishing between the yellow used for high force and the white used for excessive force. Changes to the colour scheme should be made such that the different cursor colours are clearly distinguishable from other cursor colours, and from the objects throughout the scene.

There are practical problems with introducing a system like this into a university course. The University of Glasgow Veterinary School has up to one hundred students in a year that would require to be taught using similar methods. The length of the examinations used in the experiment ranged between 121 seconds and 169 seconds. This would require a large amount of time and resources for expert assessors to assess even a small number of examinations from each student. An appropriate solution would be to look to remove the need for an assessor. By identifying common behaviours and techniques for poor examinations, it could be possible to provide the student with a self learning tool where each examination could be assessed automatically and the student could immediately receive feedback on his or her performance. However, the feedback that the student received would not be as rich as that received from a human assessor. The human assessor can direct the advice given more towards the individual students and their examinations, indicating more accurately where they went wrong, and how to correct the problem. A student can question the teacher about his or her explanation, and can therefore gain a fuller insight into the problems with his or her examination. This is discussed in more depth in Chapter 7

6.9 Conclusions

This chapter has introduced the concept of multimodal cues to augment Virtual Reality medical simulations. A description of potentially useful multimodal cues has been introduced, and a description of implementation details for each has been presented.

It has also discussed a novel experiment that has examined one form of cue as a device for allowing a teacher to assess a student. By allowing an expert to introduce errors into examinations, it was possible to test the visualisation system developed to display the graphical playback cues described.

The results discussed suggest that graphical playback cues were used successfully to allow a teacher to assess an examination. These results were demonstrated with the system with very little training provided to the teachers. This method offers the potential to allow a teacher to assess an examination and provide information to a student on what he or she did wrong, and how he or she can improve. The final section of this chapter discusses future work that can be carried out to improve the system, and discusses factors that could be used to provide a similar system that could automatically assess a student with no teacher present.

For the purposes of the third key research question, the experiment described in this chapter has demonstrated that it is possible to provide an environment in which an expert can assess the performance of a user. The success of the experiment suggests that assessment of the performance of a user on a medical simulator is possible using Virtual Reality visualisation techniques. Information such as whole path and contact point views of a user's performance allow visualisations of a procedure that would not be possible using a physical simulation.

7 Conclusions

7.1 Summary

This final chapter will provide a summary of the work carried out in this thesis, and will discuss the contributions and limitations of the work. Possible areas of future work are also discussed. The work described in thesis has followed through the design and evaluation of a training tool for large animal ovary palpation. The initial review of the literature determined the strengths and limitations of the current haptic technologies. From the literature presented in this thesis, it is clear that computer haptic exploration is more limited than haptic exploration in the real world. However, the evidence presented suggests that in particular force feedback devices are advanced enough to provide a user with a recognisable haptic representation of an object. This is an important result for the purposes of this thesis, as the work described relies on the haptic exploration of virtual objects.

A thorough review of the current medical simulator technology was presented, and problem areas were identified. One factor that can be taken from this review is the lack of research being carried out on medical simulators from a veterinary perspective. There are no examples in the literature of a Virtual Reality veterinary medical simulator (although many of the problems described apply equally to veterinary medicine as they do to human medicine). In addition, many of the simulators described have been developed in isolation from the traditional training courses available to novice clinicians. If simulators are to be accepted into courses as training tools, they must offer solutions to problems that are difficult to solve using traditional methods. Evaluation is identified as one key area that must be addressed if medical simulator training is to become widespread, and an accepted part of medical and veterinary medical courses. We must know if any simulator developed actually improves the teaching of a medical procedure. The difficulties caused by evaluation due to ethical considerations make this a non-trivial area that is still to be addressed. This is one area in which there has been a dearth of research presented. It was further argued that performance feedback was a key feature that a Virtual Reality training tool should provide.

The work described in the thesis is based around the deficiencies identified during the literature review. Initially the design of a simulator was considered, and a proven computer aided learning system design technique was employed to gather requirements for a simulator. The current equine reproductive course at Glasgow Veterinary School is described, and the potential benefits available from incorporating a simulator are highlighted. The requirements discussed in Chapter 4 indicate what simulator features would be beneficial to current teaching methods.

The simulator was designed in close collaboration with expert veterinarians. A thorough evaluation was performed, and the results described. This study incorporated three experiments that tested different properties of the simulator. The first looked at construct validity, and provided a number of measures such as performance, time, workload, and path and force data to examine differences in performance between novices and experts on the simulator. The second experiment looked at the effects of multiple training sessions on simulator performance, time taken, and workload measures. The third study built on these results to provide a comparison between traditional and Virtual Reality teaching methods for large animal ovary palpation.

Finally, multimodal cues were discussed as a potentially useful method for training and assessing a novice user. As a case study, a graphic playback cue was evaluated as a method for allowing an experienced veterinarian to assess the performance of a user.

7.2 Contribution

The work carried out was based on the weaknesses identified in the literature review. The contributions discussed have been based around the three research questions from Chapter 1.

One other important factor outwith the research questions that can be taken from the literature review is that this is the first example of a Virtual Reality medical simulator approached from a veterinary perspective. Although there are many examples of training simulators as a research area in human medicine, there have been no other examples of work from the veterinary field. However, it is clear from the interviews presented in Chapter 4 that many of the training problems experienced by the medical profession are equally prominent during veterinary training. Michell [77] emphasises this as he describes the common need for competence testing and certification in both professions. The system discussed in this thesis is the first stage of addressing this deficiency.

Question 1 of the key research questions described in Chapter 1 dealt with the integration of medical simulators into a veterinary training course. It is clear from previous literature that a failure in the design process can lead to problems with a simulator not addressing the real learning needs of a student (and being more of a demonstration of technology than a learning aid). Chapter 4 addresses the first key research question. Particular problems due to the ethical guidelines and the lack of resources available are highlighted in relation to horse ovary palpation. The work described in this thesis has demonstrated how a CAL technique can be used to identify problem areas in a veterinary medical course, and provide requirements to ensure that the simulator becomes a useful training aid. The activity charts shown in Appendix A.5 illustrate what can be achieved with the inclusion of a simulator. It is clear from these charts that a simulator can be used to provide channels of feedback to the student that would be difficult to provide using any traditional methods. Many other procedures suffer from similar problems of lack of training opportunities due to restricted resources, and difficulties in assessing novices objectively. The activity charts presented illustrate the areas in which a simulator can provide benefit.

This thesis is the first example of a haptic medical simulator that has used accepted computer aided learning techniques in order to provide goals and requirements for a simulator that would be useful in the situation it was intended to be used in. One iteration of the ABC method developed by Montgomery-Masters [79] has been followed through for this thesis. Requirements for the simulator have been developed and a study has been performed to show how a Virtual Reality palpation training tool would integrate into the veterinary course at Glasgow University Veterinary School.

Question 2 of the key research questions required that evaluation of the simulator be performed. From the literature review presented in Chapter 3, it is clear that this is an area that has been neglected by other researchers. This is an essential area that must be addressed before Virtual Reality simulators can be adopted into traditional medical and veterinary teaching practices. This thesis describes a thorough evaluation of the simulator that incorporates features that have not previously been used to evaluate medical simulators. Initially, a construct validity study was carried out. This is the most common technique presented in the literature for providing a validation for a training simulator. On top of the performance metrics described in this section, workload results were also recorded. It can be argued that this provides an equally important metric for a construct validity experiment although it has been ignored until now. Post procedure analysis of the examination has been used to determine if there are differences in behaviour of the different groups using the simulator. This method has not previously been used for assessing performance on a palpation simulator.

The second evaluation experiment examined the development of skills over time when using the palpation simulator. This represents an in-depth look at simulator performance over several training session. Again, workload data was recorded. Unlike any previous study, the retention of skills was tested using a simulator session a month after the fourth training session. This experiment shows the potential benefits that medical simulators can provide in training. A consistent improvement in performance was displayed, while time taken to perform the simulated procedure was reduced. No significant drop in performance was detected between the last two sessions spaced one month apart` showing that the skills taught were robust.

It is clear that showing improvement in simulator performance is not in itself enough to provide a validation for the simulator (the skills taught could be the wrong ones for the real life procedure). Therefore, a further study was described that tested the real world training effects of the simulator against traditional training methods. These results can be viewed as encouraging for a potential training simulation. Students who had exposure to specimen tracts found similar numbers of follicles in similar times than those who had no previous exposure to the tracts but had been trained on a simulator. No previous study had been conducted that examined the effect of multiple virtual training sessions, and directly compared simulator performance to real task performance.

These experiments represent one of the most comprehensive studies in the literature. They were designed to test different facets of the simulator to ensure that it was an effective training tool. The third evaluation was designed to build on the results of the previous study.

Further, with performance feedback being identified as a key area of medical simulators, Chapter 6 describes a set of multimodal cues that can be integrated into a Virtual Reality medical simulator to further train or assess a user. A novel evaluation study was carried out to test a visualisation and assessment system for HOPS. This study represents the first design and evaluation of such a system in the literature. By introducing deliberate errors into an examination, it was possible to show that these errors could be detected by experienced veterinarians when viewing the examination through graphical playback cues. This is the first time that such an evaluation has been conducted. This allows the third key research question to be answered positively.

7.3 Limitations

There are a number of limitations to this work that will now be addressed in this section. Only one iteration of the ABC design method has been carried out. Both the ABC method and the medical simulator design method recommended by Higgins *et al.* [44] suggest multiple iterations of the design process be carried out in order to produce a tool that provides training for the tasks required. Montgomery-Masters suggests a repeat of the design implementation and evaluation stages of the process to refine the system. With the time constraints placed on the research, further iterations of the ABC method were not possible. It can be argued that a repeat of the design, implementation and evaluation stages would not have added to the thesis from a research perspective. The study has been used to illustrate the importance of adopting an accepted design methodology when considering integration of a simulator with the current traditional methods of training. It has shown that computer aided learning methods and in particular the ABC method can prove useful in the design of a medical simulator.

One limitation of the design of the simulator is that the user's interactions are restricted to the one point of contact available from the PHANToM. The PHANToM provides good quality feedback to the user through this one point, but restricts users from grasping objects within the scene. This has previously been discussed as a problem faced by the expert group in Section 5.2.6 when placing follicles on an ovary. Significant hardware developments must occur before a more suitable device can be considered for the task. Results suggest however that while one point of contact limits interaction with the simulator, it was still possible to provide a useful training environment for the students. One possible currently available solution to this problem is discussed in Section 7.4.1.

Finally the lack of a second control group must be considered a limitation of the comparison experiment between virtual and traditional training methods. This group would have received neither the virtual training nor the traditional anatomy lab training. Since significant differences were not discovered between the performance of the two groups, it is difficult to identify how much skill transfer each of the methods confers on the students. A second control group would have provided a means to identify this. When conducting these training experiments, it is essential to use a method that will not negatively affect the amount of traditionally accepted training that the participants in the study receive. It can therefore be difficult to conduct these experiments without interfering with the traditional training courses. The study described still examines a valid measure. Anatomy laboratory training is currently the best feasible method of providing training to veterinary students early in their course. These laboratory tutorial sessions have been included in the veterinary course (as well as in human medicine) as an accepted method of training for decades. It is encouraging to note similarities in the results displayed by both the traditionally trained and the virtually trained groups. The benefit of the virtual training is that other features (such as multimodal cues) can be included to aid learning further and these are not possible with the traditional methods.

7.4 Future Work

7.4.1 Increasing the Number of Contact Points

One factor that limits the fidelity of a simulation is the method of interaction. In Chapter 2, the benefits and limitations of the PHANToM force feedback device are discussed. A major limitation of this device is that the user is restricted to interacting with the environment with a single point of contact. However, the use of multiple PHANToMs allows more than one PHANToM cursor to be included in the environment. Each of these cursors will still represent a single infinitely small point in the environment, however, the user will now be able to touch a virtual object at more than one point at once.

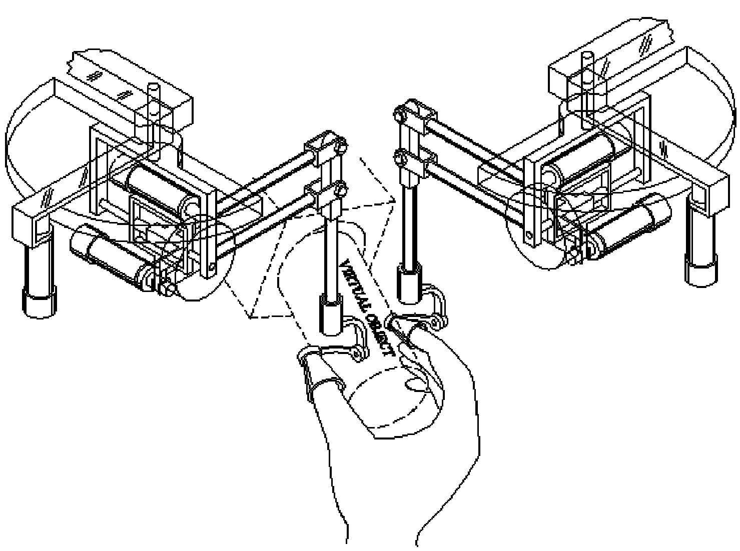


Figure 57. A diagrammatic representation (from SensAble Technologies [103]) illustrating a user interacting with a virtual object using two PHANToM devices.

Using two PHANToM devices - attached to a finger and thumb on the same hand as shown in Figure 57 - gripping of objects is possible. With one point of contact it is difficult to palpate a moveable object, as the user cannot support the object to apply a reaction force to the palpation. With two points of contact, one can be used to palpate the moveable object and the other can prevent the object from moving by providing the reaction force. It can be argued that increasing the number of contact points with which the user can explore an object – and therefore increasing the potential information available to the user at any one time – may lead to faster and more accurate object identification. This is not necessarily the case. The examination of the object is still temporal and the user must attempt to integrate the separate points of contact to form a model of the object being explored. Studies have shown it possible to interact successfully with an environment using two PHANToM devices, however, there have been no studies to show that it provides benefit in object identification. This would be an important factor on whether inclusion of this method in a simulator will provide any benefits. A pilot study conducted in the HOPS environment with two PHANToM devices of different sizes attached to a participant's thumb and forefinger suggested that providing two points of contact would not necessarily provide a method of locating and sizing follicles more accurately and more quickly. A small number of participants and calibration problems due to the different device sizes, however, may have contributed to the difficulty of the two PHANToM condition. Further work with two similar sized PHANToMs is required.

In the case above, the second device is used to provide an extra point of contact with a virtual object. Alternatively, the extra interaction point could be used to control the palpated object. In a bovine ovary palpation procedure, the veterinarian will generally cup the ovary in his or her hand and palpate the ovary with the thumb. Allowing one finger to control the movement of a virtual ovary may provide a more realistic environment for the examination and provide the user with a better awareness of his or her cursor position on the ovary. The finger controlling the virtual object could feel the reaction force from the palpating cursor, and provide the reaction force to stop the object from moving. However, there is no literature to suggest that object identification using a force feedback device will be successful using this method. This method may provide an environment that allows interactions that are closer to the interactions in a real life palpation examination, but if it increases the difficulty of the identification of surface features, then this may provide a confounding factor to the training. Users may be training to adapt to the difficulty of the system instead of training in the skills involved in ovary palpation.

7.4.2 Examination Assessment Through Recorded Path and Force Data

Chapter 6 described an experiment that was successful in showing that an environment for viewing a recorded examination can be used to assess the performance of a user. It was also argued that providing performance feedback to users as they practice on the simulator is an important part of learning. Chapter 3 argued that certification is another important area where simulators would be beneficial to provide a means of assessing a clinician's performance. It was also noted however that for the five examinations, it took most participants well over thirty minutes. Although the time taken to assess an examination would be expected to decrease as the assessor became used to the system, the time required would still be unfeasibly long for a university cohort of approximately 100 students.

The next stage of developing an environment for assessment would be to use the results gathered from this experiment to develop an automatic computer based system that would provide feedback to assess the performance of a user without the need for an expert teacher. This will provide the basis for a system to provide a training environment for students without the need for a teacher to be present. One of the major benefits of such a system would be to provide vastly increased access to interested students outwith their university course. A method of measuring each of the factors identified during the experiment in Chapter 6 must be developed. For factors such as too much force being used or ovary coverage, this could be straightforward when given path and force information. Measuring the number of correctly diagnosed follicles would be possible by allowing the user to enter follicle details into the system. Factors such as search strategy or overall rating would be more difficult to calculate. Further research would be required before a system could be implemented that incorporated these factors.

7.4.3 The Evaluation of Multimodal Cues to Provide Training

Chapter 6 introduced several examples of multimodal cues and how they could be developed to provide performance feedback to a user. Feygin [36] demonstrated that pre-recorded graphic and haptic cues can be used to train a user to follow three-dimensional gestures. For a medical simulator however, there will inevitably be contact with virtual objects. The reaction force from these objects will directly affect the movement of the user. A future experiment should be conducted to evaluate the effectiveness of different forms of multimodal for training. The experiment could be designed to examine the effectiveness of providing training using three different combinations of multimodal cues. Participants would be asked to experience a gesture, and then attempt to follow the path of the gesture without the cue. The first condition would use pre-recorded graphical cues only to present the user with the gesture. In the second condition, the participant would feel a pre-recorded haptic gesture - as well as the visual cue - that allows him or her to follow the gesture, and then try to recreate it without the cue. The final condition would use a pre-recorded graphical cue combined with an audio cue to provide force information. Measurements could be made in each case in the accuracy of the recreation of the gesture, and the accuracy of the force used (estimated as the reaction force from the object). This study would attempt to demonstrate the usefulness of the multimodal cues in particular for the early stages of training. Different cues that provide different levels of assistance to a variety of skill levels could then be evaluated in subsequent studies. For example, a student towards the end of his or her training may only need audio guidance, where as a novice may benefit more from direct haptic guidance or a combination of cues. Again it must be noted that further studies would be needed to show the skills learned using these methods transferred over to the real life procedure

7.5 Conclusions

In conclusion, this thesis has followed through the design, implementation and evaluation of a novel Virtual Reality veterinary palpation simulator. It has brought together and advanced work from a diverse set of fields including psychology, human-computer interaction, computer aided learning, medicine, and veterinary medicine and has set out to answer three key research questions. These involve the design and integration of a veterinary simulator into a training course, the evaluation of training benefits provided by the simulator and the possibility of assessing the performance of a user when interacting with the simulator. This thesis has been able to provide a strong justification for answering positively to the first and third research questions. The second research question was investigated thoroughly. However, despite encouraging similarities in the performance between virtual and traditional training methods, no significant results were achieved to provide a strong basis for answering this question positively. At every stage during this project, evaluation of the work has been included as a key aspect. By this method, the success or failure of different stages of the work can be assessed. This is an essential factor in the development of medical simulators if they are to prove their benefits to the medical and veterinary medical world. The work described was able to demonstrate a novel design and evaluation of a Virtual Reality palpation simulator within the constraints of a real teaching environment.

Appendix A: Requirements Data Gathered Through Interviews

A.1 Introduction

This Appendix contains the data gathered through the requirement gathering techniques described in Chapter 4 and the analysis of the data. Transcripts of two interviews, and notes gathered from a third interview are presented in A.2. The classification of the interview statements is presented in A.3. A.4 contains a table showing the differences in classification on sample text from each of the reviewers. Section A.5 shows the Activity Charts built using data gathered through the interviews.

A.2 Interviews

Interview 1 Transcript

The interviewer (Andrew Crossan) is A. All sections that were unclear from the tape are marked with the symbols ****.

A: It's just generally about course matters and what students have to do...

S: Are we talking specifically about Glasgow students or students from other unis

A: Most of the stuff is about Glasgow, but I know Monika mentioned that Edinburgh got a big horse department as well and things are done differently there.

S: Yes well different vet schools have different approaches. London, Liverpool, Cambridge... when it boils down to it everyone other than us really... Bristol have little herds of ponies that they use maybe in more... so their students probably get slightly more experience than our students do. I mean I can tell you what the ideal is or... like in a way, you can just ask the questions and provided it doesn't use up all your tape.

A: There's a good 4 hours to go so... My brother actually went to Bristol and what he was saying was that he's never actually rectalled a horse because they weren't allowed to do it there.

S: Funnily enough, I worked there quite a long time ago now, but at that time, they did have a rather legendary guy called John David who taught that stuff and at that time they did have a little herd of ponies that they used for that, but bottom line is that there are really 2 issues that arise from that. Now one is that its very costly to do that and having a little herd of ponies is almost worse than having none because we've got so many students that there's nothing worse with students than when one gets the chance to do it, but another one doesn't they feel madly cheated and so on and so forth, and so from that point of view, I guess the logistical things are the mechanistic logics are.... but the other thing is that there is a kind of ethical issue that does arise. Is it fair to abuse the pony's backside basically and all in the interests of vet students? So your brother's a vet now is he?

A: He's working in a small animal practice down in Burnley. He's worked down there for 2 years now and he's quite enjoyed it. I think he'd like to do some large animal stuff as well but I think at the moment he's quite glad that he's not what with foot and mouth.

S: At the moment yes...

A: So there's 13 questions. Some of them overlap, and really just about the course, so what the course objectives from first year to final year, so what should students be able to do at the end of each year.

S: Well currently, and this is all up for change in fact it is changing literally this year as far as the final year students are concerned. But currently, all the students are expected exactly the same course, although all that's going to change. So I'll come to that in a minute... As you probably know, there is at the moment as you probably know a progressive course such that first and second year subjects that would relate to this subject are specifically anatomy of equine reproductive tract, male and female, and physiology and it would be true to say that the physiology is largely taught as a kind of generic thing so that they wouldn't have anything in second year at least specific about the mare, other than the lecture material people may use to illustrate the difference between cat and mare and whatever because there are fundamental differences as you've probably at least vaguely surmised.

A: Well recently I sent Dominic Mellor up some modelling clay, and he sent me back a model of a horse and a cow ovary and I was just surprised that the horse model was about that size, and the cow ovary was a wee tiny thing.

S: So you know, there are fundamental differences you can compare and contrast, but it would certainly be at a pretty superficial level, and the objectives would be more you know to know the overall hormonal events that control and fundamental things about anatomy, as a basis for when it comes to do ovarian palpation. So it'll be knowing structures, knowing functions really, but not at any great detailed level. The only third year subject that would even vaguely encroach and again it would be pretty minimalistic would be the methods of manipulating the reproductive cycle i.e. with various administered medications, so how you could induce ovulation or how you could prevent esterous behaviour at all... that kind of thing. And again they will often be delivered as "these are the induced prostoglands into you can use progesterones or whatever" and they would be as far as the mares concerned **** mares and pigs you often give oral cogesterogens because that's just the way you do it but it would be kind of at a level no vet. Fourth year would really be the first time when any - and again this is Glasgow fourth year it would be different elsewhere - essentially when we would get into a little bit more depth and currently there are 8 soon to be reduced to 4 lectures on equine reproduction. So that includes male and female, mostly female. And so that would then include a kind of hour of kind of normal cycles and manipulation of cycles and then an hour of kind of fertility practices, you know ways of maximising things, and ways of manipulating ovulation. Using ultrasound and so on. And now as far as ovarian palpation is concerned in the mare and that does come in to that, because the likeliest thing you would be doing is sequential monitoring the physical characteristics of ovaries to predict when the mare is near to ovulation, cause ideally and... in the ideal world, you really only want to serve the mare once and that is essentially because every time the stallion enters the mare, he contaminates... the mare is very prone to endemitritous. So fertility is likely to be better if there is only one mating, but to make sure there's a good chance of pregnancy, you want it to be when she's in ovulation. The old fashion way of doing it was that the mare as soon as she was in standing esters would be mated every second day until she was no longer keen on the idea. And that might need five matings, which is fundamentally a bad thing to do. So the purpose of predicting ovulation, and that's probably the most important that would have to know. Sorry, students would have to learn... the important thing vets need to know is to be the person that says this mare is going to ovulate within the next 24hours, therefore mate her today. Or inseminate her, which could be done by artificial insemination. That is probably the most important objective in terms of the physical skill, because it is really a physical skill that this would help them to do. Now where there is an irony for you I think I don't want to shatter your dreams too much that now days, that's nearly done exclusively by ultrasound. I wouldn't lose heart with that because I think it's a much easier proof of concept. I mean its a long time since I used the model, but proof of concept that you can recreate at least the general sensations and develop the kind of skills means that it could be I think transferred… it would be relatively easy to adapt in order to teach people other things, and nothing to do with reproductive tracts, in particular diagnosing different types of colic. You could do that because overall the physical methodology and the fact that you've shown that you can use it to appreciate size and textures and that sort of thing... softness, turgidity, that as a proof of concept is good. So it would probably be the correct thing to start with. Its different in a cow, I mean there, although they do use ultrasound, its more for pregnancy diagnosis there, but they still feel they're ovaries. Cows get more ovarian disorders than horses probably, or they're more common in cows. There are more cows. I'm not sure if it's true to say that they're more common, but certainly, vets will encounter more. So detecting things like cystic ovaries which is a thing that's done in cows is practised frequently. Cystic ovaries don't exist in the horse. So in fourth year, the objectives of fourth year are really to make people appreciate the main focus of fertility management in the mare because bottom line is that's for reproduction in the mare comes down to. Its actually a relatively small number of people ever will do it to some extent. What it comes down to is 2 almost discreet areas of it, or maybe 3. Manipulation of the ovarian cycle, and prediction of ovulation to optimise the time of insemination - so mating, that's what we're just been talking about. Number 2, early pregnancy diagnosis. Which again tends to be done by ultrasound, very early - 15... 16... 19 days. And then the third main component of equine reproduction is investigating reasons of infertility, in other words, why is it not getting pregnant. And those are far and away.. those would be the 3 main components, and that's probably what the objective of the fourth year course is. To get over those things. There are then a short catalogue of specific diseases and things that the students would be aware of. I mean things like retained placenta for example. You could learn in the cow that that's not a terribly big deal, but in horse, its a life threatening situation probably the biggest emergency in equine practice bar bleeding to death. So basically, they learn about these 3 things, and then there's a sort of catalogue of... currently there's no hands on skills. That is not to say that access to what you do is not... in fact it would definitely be a very meaningful thing to do in the context of a lab that supported or was complementary to that group of lectures. Probably to use examples of extremes. Here is a non-cycling mare with no ovarian structures. Here's your mare on the point of ovulation. But you would be needing to teach all of the students... or I don't think it would be appropriate that

100 students would need to go and know those skills in a way they would have to go and apply them. I mean your brothers a good example. He is never going to do that... unless he finds a partner that happens to have a horse, which lots of vets do and then the partner says is my mare pregnant. In other words, if 80% of vets end up in small animal practices what is the point of teaching them those skills. Well the point is to illustrate a concept. Literally to illustrate physiological events to make them graphic, but not to teach them skills. Currently in the final year then is get all the students experience of palpating ovaries and uterus in mares. Realistically, I think we nearly achieved that. Most students will have had they're hand in several mares and that would be to try to get them to learn some physical skills. As it so happens, this next year, this is the first year we're going to be having species tracking so not all students will come onto the equine course. So we'll have a smaller group of them and theoretically at least, they will be the ones that will really want to do horses. So then I think we would definitely make a more determined effort to try and make sure… in other words they probably are core skills to that group of people when we still have at the moment 100 students coming through… most of them have to come through the equine rotation even though they're never really going to use any of the species specific stuff that they get there.

A: Are your numbers growing or shrinking

S: They're growing but now capped. I mean basically we gone in a sort of… well certainly in a 10 year thing from having 60 to having 100 and numbers do make these things very much much more difficult to be able to give them all experience. Particularly again when many of the students are the people with the biggest ethical concerns. This faculty does have believe it or not and this causes a certain amount of hilarity a code of practice about how many times a rectum can be used a week and that applies to cows and things. If you go back historically – this is general background waffle – historically probably the second year students as part of their anatomy thing use to get – and there was big excitement to go and do your James Herriot bit – every second year student went down and brutalised some cows backsides… and some of the students started complaining about this saying it wasn't fair on the cows… and possibly it wasn't. It certainly becomes less fair when you've got more and more students and fewer and fewer cows.

A: They let me go along with a fourth or final year tutorial and that was lots of fun. The farm up the top of the road had maybe 10 cows in stalls and 10 students, but they were very strict about how many times you got to put your hand… they just stood there anyway though.

S:I think about it… overall I suppose what I would say about in terms of an objective of the teaching at undergraduate level, the objective or expectations of their skill is pretty limited. I mean I wouldn't really expect the students to be anything like as competent in this specific subject as I would expect them to be in lameness examination, because even within the very specific sort of horse vets, huge numbers of them do not do this. Probably, I would think you could say that 80% of all mare fertility reproductive work is probably done by 100 people in the whole country… well maybe 200. It's a very small number of people who are actually doing this as a living. I mean it is almost at a conceptual level, and if they're going to become really skilled in it then it's a postgraduate subject for them. I think that should be good news for you though, because I think that means that the absolute recreation of real life sensations in the model isn't massively important. We're not going to try to train out and out – or I wouldn't but you might – I wouldn't have thought that your trying to think that this is absolutely real life you've recreated, and that really exactly how it is. Its kind of like this is sort of what it will be like, and this is scenario A, B, C, D, E, and F, and if they do it often enough, they will get the general feel.

A: A Lot of the stuff we've talked about… we know certainly just now it's not possible to get anything approaching realism. The idea is to get across the concept, and the general feel of as close an approximation as we can. So we've talked about various things we might do. We've talked about maybe switching to other examinations later. Someone mentioned cataract surgery in dogs which is apparently very rare… maybe a few experts in Scotland can do it and that's about it. But I think Stuart and myself are quite keen to carry this through… as you say, use the proof of concept to lead onto other things. But some of the things are tutorial ideas where you can… have you seen the device?

S: yes, I've used it. It really cuts your finger off.

A:I had a subject in for an hour long experiment the other day, and they had to stop after half an hour because their hand was turning red. You can also use that also for recording movements, and also for driving users through a series of motions. Something we've been talking about is in augmenting the simulation. Rather than improving the models, augmenting the simulation to try and provide things that you couldn't get in real life. So maybe use a recording of a student's examination to assess their performance, so maybe they missed this bit, or they made this incision too long. The other thing that we're think of that's slightly harder is – I don't know whether its going to come off just yet – is actually recording a vets movements and playing it back to a students so that it kind of drags them through the motions, and we've got something working roughly that will drag you… its still quite rough.

S: When Stuart was a student and I, there was a somewhat eccentric or idiosyncratic at least woman who use to teach us reproduction in all species, and she's retired now and she use to… don't put this in any report… she use to particularly in cattle… she's really enthusiastic she tried to get students to feel the right things. And eventually - she couldn't believe they were so incompetent - that she would grab their hand with her hand put it in the cow, and drag you literally to the spot. So I suppose this is slightly more sophisticated, and more pleasant for the cow. Certainly, I mean Anthony (Clement) was very enthusiastic after the thing yesterday (parameter setting experiment)… well maybe there's a kind of novelty value, and I think maybe with someone like me whose a real computer dinosaur and Stuart sent me up to see Aiden (Glendye) years ago, and I suppose your just sort of – if you know nothing of these things – it is just sort of amazing. Certainly there would be, although it may not be a useful way to focus and spend your time. Within the horse, rectal examination for diagnosis of certain kinds of… I don't know if you know what colic means but they're intestinal disorders that cause abdominal pains. It always forever be – no matter what new technologies come along – it will always be an essential part of a colic examination. And that is a very common scenario that anyone who ends up doing any horse practice will have to do, and actually because it's a naturally occurring disease no matter supposing you have 100 pony herd your not going to get many cases of colic. And because the horses that come here – I don't know if Antony showed you round its not quite finished yet - they all come in and they're owned by people so you can't really turn a student loose on them. So you can't give the students any sort of experience. I mean you could show them slides of the… I could show you slides of the pathology that they're trying to feel per rectum. So what they're trying to do is workout has it got distended loops of intestine, and are they small intestinal loops… are they large intestines… are they distended with gas…are they distended with fluid… are they distended with ingesta. That is exactly the questions they go through. It's a simple iterative process. But I mean Antony was completely convinced - and it may well be that your model is much more sophisticated than whenever I last used it – but he was absolutely convinced that you could use that again to produce really useful teaching aids. And that would be really useful, and it would be massively fundable. The only thing about it is certainly… and its funny that Antony and Rob – who were two of the people that used it yesterday that they sort of kind of felt that a whole hand in that involved your whole hand, shoulder movements… they felt and they may be wrong of course… they felt that that would make it more realistic. That depends whether you feel realistic is important because in a way its kind of can you tell the difference between small intestine distension and large intestine… can you tell the difference between gas distension and solid distension because they are fundamental questions in to making the diagnosis. They're absolutely the things we bang on and on about to the students because we feel that's what they should be asking themselves. They don't have a tick sheet… they should be subconscious. It's exactly what you've proved to my mind at least that you can do with this. You can get something… and OK those wouldn't be spherical structures, they would be tubular structures. For example, probably not today, but maybe today, I could get the slides out I could get you images… but small intestinal distension inside a horse feels for all the world as if someone's got a whole load of… the thickness of your wrist… they feel like distended bicycle tubes. You feel as if you're running your hand over a whole load of bicycle tubes.

A: How do you tell the difference between say a gas-distended intestine and a solid distended intestine?

S: Because in the solid distended intestine there would be sort of less give, and gassy would be sort of an amorphous thing out of Doctor Who. SO there is a different texture and a different level of resistance really. I'm sure you… well the level I… when I tried to help Aiden a little bit… what's ever happened to Aiden.

A: He's working at Unisys in Glasgow. He's been out working in Holland.

S: Anyway this was something I was sure could be done, and it would be if you wanted to develop that aspect of it… it would be…. And although it would be really useful for students, it would be useful for vets who are doing professional update courses and that sort of thing.

A: I'm really the only person working on medical haptics, but there are a few other people working with other things, but Steve my supervisor is always looking for new ideas and new ways of getting funding and that sort of thing. We would like to expand a bit because its frustrating that we've got all these ideas but just not enough time to do them. Its certainly something especially if its something that employs give and texture, its something that we could provide… we can certainly provide different situations from that point of view for these different variables.

S: Give texture and shape. I'm presuming shape is possible.

A: Shape will be possible. The models are still quite similar to the ones that you used, but Dominic recently made up clay models that were scanned in at our graphics lab, and imported into the simulation. I'm not sure if we'll use them yet because there are still a few problems with them. They made of basically triangles stuck together, so its not smooth spheres as it was before… although you can't generally see them your sense of touch seems to be a lot more sensitive to these sort of things…. So I'm not quite sure if I'll use these yet. The nice thing is we've got a very well funded graphics lab that have many 3d scanners that allow you to build physical models then scan them into computer and manipulate them as a virtual object. And that's the sort of thing that if we had a few ideas and could get some people interested from the graphics lab, we could really kind of…

S: Well what we have just acquired at not inconsiderable expense – Antony maybe mentioned we were getting a physical phantom horse that was suppose to… which is a fibreglass back end of a horse with several sets of intestines that you can stick inside it in different constructions. The intestines are basically inflatable, so I'm sure we could bring them along to the graphics lab and do it. The only thing is they're inflatable with gas obviously so you'd need to start trying to get the turgidity of fluid distension by different means… I'm sure we could achieve it.

A: One of the problems not with shape… you can scan in the shape perfectly, but there's no way of scanning in your haptic properties. You have to come up with them yourself, hence the fact I was up here before getting people to mess about with properties. I'd be very interested in seeing the horse.

S: Antony can show you it… even just for some amusing photographs that you can put in your thesis.

A: We thought about trying to get one of a cross section that we could just kind of put in the background of the simulator. We never quite got round to it… so that was question 1. Question 2 is how each of these objectives are met… basically going through lectures and…

S: At the moment it would be true to say that first second and third year are all lecture based, and nothing else. Fourth is lecture based… so first through to fourth year is lecture based purely. The lectures… nearly all vet student lectures will be reinforced with visual materials i.e. slides of specimens or ultrasound videos and that sort of stuff… and in fifth year final year there are practicals i.e. a horse back side and a demonstrator reinforced by simultaneous ultrasound and palpation. And we have - which your welcome to take away if you want – a couple of CD ROMS on equine reproduction and a video on early pregnancy diagnosis.

A: Yes that would be good. So the next would be how well could each of these course objectives be met. You were saying before its not so much the practical skill, but…

S: Because - if you like – because me and its only me, other people might take a different view – but my view is the objectives are of an awareness at a bit more than a conceptual level, but not massively more, which I think we achieve. I mean we don't set ourselves very ambitious objectives probably so we do achieve them. But the biggest limitation is probably the group numbers and as I said we've actually got something in hand to try and effectively try and reduce the number of students that want to do a practical.

A: So what would you say the core skills for the procedure really are?

S: By core skills… the skill is being able to identify ovaries, identify structures: ovaries and uterus. Appreciate uterine edima, uterine distension. That's probably it. Those are very specific, but do you mean what kind of things would make somebody good or bad at this.

A: Well what would you say the hardest things to teach are?

S: That's quite interesting… The hardest thing to teach probably is early pregnancy diagnosis… so that's a uterine event obviously. That's probably harder to teach than has this horse got developing follicles. Is the horse near to ovulation? Has the horse ovulated? Probably overall that is…

A: So what sort of things are you looking for for early pregnancy diagnosis?

S: Well as practised in the horse, how is it done or…

A: How is it done really…

S: Well its essentially done by ultrasound and the reason its maybe difficult is that there are conditions of the uterine wall that could mimic pregnancy in particular uterine cysts. Because all and early pregnancy looks lie is a kind of Safeway's tomato sized fluid filled lump, so it looks like a black circle about that size. So it looks like that. That is probably the most… I think it's more difficult… The other thing is that I think it's a psychological thing rather than a physical thing. If you say yes it's pregnant of no it isn't, it affects the immediate veterinary management. So the repercussions if you get it wrong are probably greater. It's more difficult to be confident about it. Probably a psychological thing rather than a physical thing. I don't know, I've never tested it. It's just my belief that it's probably a little more difficult.

A: So how would you go about differentiating between a cyst and a…

S: Well the clever thing you would do is map its uterus before you started breeding so you would know if it happens to have these cysts so that when you now find them three weeks after it had been inseminated, you would know that they were there all along, and they were not the pregnancy. The other thing is you would repeat the ultrasound at an interval of several days, and the ultrasound will progress, the cysts won't.

A: So is training provided from sources outside the department…

S: Well 2 things. One of the fourth year lectures are provided by an outside source… i.e. I guy who makes his living doing this and nothing else. That is a guy who makes his living doing equine reproduction work and nothing else. And he comes on a hired gun basis and gives the student lectures. And he's highly experienced. He used to come from and academic job in Holland. In terms of the handful of students going on placements, and it is literally that… maybe 5% of students who really want to get into equine reproduction in a major way. We would help them get… I'm sure you've heard from our brother about going to see practice and all that… We would get them placements in practices where this is a major part of the practice activity. And this will quite often involve getting them to go overseas to Utrecht, Florida, Texas.

A: It's a dirty job but someone's got to do it.

S: Exactly… New Zealand. Obviously there are reasons why they go to the likes of these places. But anyway that is the way to get them additional experience.

A: Outside placements, what part does self-learning play in the curriculum?

S: Well essentially probably self-learning would probably be the relatively few who are interested would make use of the CD ROM type stuff to had. And to be honest, its commercially produced and purchased, and its not of sensational quality.

A: Is this anything like CLIVE

S: No it's nothing like that. Essentially all it is, is digitisation of a course which is run for postgraduate level for vets in practice, and its basically digitisation of some lectures, some of which are very boring. Many of which are very boring.

A: So how is time divided up with lectures labs and practical experience? I think you've gone some way to answering that already.

S: Yes, Well overall, it doesn't comprise a very large component of the equine teaching. I couldn't give you an exact figure. But the student hours on the subject year 1 through to 5 would probably be about… 30 now that is out of something in the order of believe it or not 2800 hours.

A: So presumably that's similar to a lot of the courses down the road.

S: I would say that we're probably… that would be less than most other places. I suspect its probably similar to Bristol, but probably less than some of the other places… it might be at the maximum it would be twice that so its still pretty small…. Less than 2%. Put that into… its almost a meaningless statistic without... I could find out for you maybe there will be other important subjects – clinical subjects – in a way, they're probably… there are 25 to 40 clinical disciplines. So in a way it's not surprising. In equine lameness for example, the hours would be in that probably – could be 100… maybe 120. But that would be about right because its massively more important subject. It's probably relative to its own importance. That figure of 30 hours for these students who are… we're calling it sort of equine tracking students, that figure's going to go up for them. So the tracking is only going to apply to final year but I would say that that figure will probably only go to 50 hours. For them, they are going to get massively more. ****

A: Are there any areas you'd like to see included in the course that aren't already there?

S: Yes. One thing we should be doing more really is in two things. One is on a sort of getting their level of skills at basic pre-breeding assessment and or fertility assessment of mare so yes we would for a certain proportion of students, getting them more competent in those… what are everyday procedures if you take this up as a career. The other aspect would be all to do with artificial insemination. Which isn't totally unrelated to this subject, because the key point about all this importing frozen semen from Canada at huge expense – well chilled not usually frozen because in cows, its all frozen, and it lasts for year, and you can fly it round the world and all that sort of stuff quite literally. In horses, the synchronisation of the whole thing becomes massively important – in other words, when's it going to ovulate. Artificial insemination is becoming a lot more common in horses than it was. In cows, its been going on for ever.

A:I know Antony mentioned importing expensive horse semen from Germany, and if it doesn't work, you've wasted however much money.

S: A lot of money, so we should be doing much with AI.

A: So is that relatively common now.

S: Yes it is. What is a concern, because its exactly in cows, sometimes farmers do their own insemination or its done by a technician, and in horses now, its more and more being done by technicians, and its true to say that the technicians are probably… at the point of graduation, a technician who would have no qualifications essentially at all is likely to be better than a vet. And it is quite common.

A: So, I suppose, leading on from that, what areas do think could be improved on in the course. Where could you offer more training?

S: In a way, that's what we sort of achieve. Currently each of the final year vet students comes here in equine hospital as part of their final year, which is essentially all… in the final year, there's no lectures, it's all clinical hospital training. Each student currently comes for 2 weeks. Next year the 40% who have chosen to do large animals… sorry, who have chosen to do horses as opposed to all large animals, are going to do 8 weeks so basically, we're going to have more time to teach them these physical skills. So that is the main way we can improve it. By providing time and facilities. Now we can't provide 100 horses, or 50 horses, or 40, so having teaching aids would be useful.

A: So if you did have your ideal horse ovary simulator, what would you have incorporated into it

S:I think if it was to be absolutely ideal, I suppose you would have every scenario of ovarian physical events that you could have, so that matter ovarian and uterine… Sorry was the question ovarian simulator or…

A: Any…

S: When I say every scenario, there are not that many, and they may well all be on there. But you would have everything from small hard ovaries with no follicles through to large twin ovulations, and then probably the odd scenario such as the relatively rare disease states of the ovary such as ovarian tumours. And certainly, would as ideal have it sort of various stages of pregnancy, particular early pregnancy with appropriate ovarian events that go along with that. So in other words, correlating early pregnancy, and in the same simulation, having ovaries that would match with that pregnancy.

A: One of the things again that we have talked about is if you maybe have a slider bar at the side, and pick a stage… so as you move it through, the follicles grow, then ovulation and straight through to pregnancy. So its something we have talked about, but it's a bit in the future. First thing we'd like to do is include more in the environment. So, just now, its just 2 ovaries.

S: Funnily enough, the other thing that I was just going to add is what you would definitely have is.. you'd perhaps… probably the likeliest thing that someone would confuse with the things is a ball of faeces and so if you like, having you know, a great… it probably feels like apples floating in water and which ones the ovary. And everyone's grappling around not sure whether they've got a ball of faeces or an ovary. So that would certainly be true. Certainly, if you wanted to take it to its real extreme… maybe the mare's straining on your arm so you can't actually… that's a slightly different thing. You couldn't with your current thing make that.

A: As far as we got was thinking about a wooden sheet with a hole in it, and maybe a bit of rubber tubing so that you're restricted in your motions. I think that's with current technologies as good as… I don't think they're bringing out a horse backside haptic device.

S: So in a way, those are…. Again it depends on whether you're trying to use it as a teaching aid… so as you say if you can slide it along it kind of is progressive, and you can feel in your very fingers… kind of if you slide it to 72 hours before ovulation and then as you slide it closer to ovulation, its changing in front of your hand. I'd imagine that would be a very useful thing. And you could therefore… presumably, you could set up self-assessment once they've done the training program a few times or whatever. And that from an educational point of view, that would be great, but certainly in the simulation thing, what could be there that could be confused with ovaries. What could be confused with pregnancy… a fluid filled bladder an infection within the uterus might feel like it. So I guess there are quite a few…

A: It's really that sort of flexibility…

S: Because at the moment its in a vacuum and when you put all the other bits – the abdominal organs in – then I guess if you like that that would make it more real, but I'm not sure that that's maybe.

A: I suppose it doesn't even have to be the same simulation. You might have one simulation for surface features. One of the things I've found difficult when incorporating other objects into the environment is it's really, really difficult to get a 3D spatial representation of what it should be like. It's hard to get your head round the fact that… or head round where all the organs are.

S: Yes… Something that might be, and I don't know if its happened yet there was a guy, and its not entirely representative. There was a guy who had a project going on in the States. As far as I know, he was using MRI to kind of create 3D pictures of a horse but because of in MRI scanner you can only get relatively small structures into the MRI scanner, he was actually putting a whole horse in, but they were miniature horses. I mean there are various breeds of horse which are literally that size. It wasn't like a proper big horse. Now there's probably some assumptions… probably wrong assumptions that the anatomy of a miniature horse will be exactly the same as a proper horse, because almost by definition they've got little deformed legs. Never the less, if that project is available… I don't even know the guys name, but whatever. A 3D horse, or MRI horse… you'll get a lot of MRI horse but it'll be clinical things.

A: I know in Edinburgh they've done a similar thing with a dog. Its called Lucky the Virtual Dog… its not very lucky because it's a dog cadaver, but they scanned it in using an MRI scanner, and built up a 3d graphical model.

S: There was supposed to be one for horses. I don't know whether this would help this spatial 3D thing or not but I mean I guess the question was what would be the ideal, well the ideal would be that there was ultimately at the top level, if you like something that included these other confounding things. But that might be for final year or even post grad people where as second year students could learn a lot from your slider and ovarian development. So I think things would be slightly different for different groups.

A: So how would you say if you had something like this… how would it fit into the course?

S: Well I think you would use it… or could use it probably as a supplement to second year. I.e. ovarian follicular development probably in particular and also… and again because your simulation at ovaries at the moment, I wouldn't want to hinder that. Follicular development and pregnancy diagnosis… particularly early pregnancy diagnosis could be incorporated into that and that would be second year type subjects. You could use maybe similar materials, but maybe a little bit more advanced and correlate them particularly with ultra-sonography. I think that would be really neat, and if there was a side by side a sort of as you put your hand on this thing that's developing as you slide the scale along the ultrasound appearance that would correlate with that going on side by side… would be neat. And in final year, I think you would use it as a learning tool to try and say… I think you probably would create scenarios of problem solving. This is a mare that's been mated four times and she isn't pregnant. Part of that would be ovarian assessment, and you could tie it in with that, but there would be other aspects of the investigation such as say uterine biopsy that you know… so it would be a sequential list of questions that people should ask themselves, and one of the things that **** it would be an obvious thing to do. **** their cycle. But I think it would be in a kind of problem solving tutorial.

A: I certainly like the idea of the combination of ultra-sonography. We did this thing for a medical emergency conference down in Manchester. We were basically just introducing the idea of Virtual Reality in medicine. But one of the things we did, we just did a simple pulse demo its not even an arm, its just a cube that's meant to represent this bit of the arm. You move your hand over it and if you move to a certain point, you can pick up a pulse and you can change different features about the pulse, making it stronger or faster or all sorts of things. One of the ideas we had was to combine it with a stethoscope sound… without knowing anything about it, combining it with an ECG trace.

S: Yes, I think…. It seems to me that combining something that's a physical sensation with a visual… to my mind that would be… just intuitively, hat would reach people a lot.

A: Or maybe a smell device to complete the experience…

S: I thought you were going to say taste…

A: There's actually a few really scare looking olfactory devices – which is sense of smell –which looks like big claws stuck up peoples nose. They do exist, but there's not many taste ones about. And it's maybe not so useful in horse ovary palpation.

S: Not to my knowledge anyway.

A: The last thing really is can you think of anything else that I might have missed or…

S: I wouldn't think so. I don't know if this is something that you've missed, but it might be worth kind of going through a few of the questions, maybe just a few test drives with some students. I don't know if you've thought of that. I presume you've seen Aiden's dissertation. I think what he was trying to do is to try to get students with different experiences or whatever and show that it wasn't worse than the real thing.

A: I'm actually working on something similar just now. Hopefully Stuart's lined up some students for when they come back from Easter.

S: I can guess which ones he's lined up as well.

A: Very similar to Aiden's but the idea being… I'm not sure if I've got it in here somewhere. There's 3 groups. We have a group with… well two groups with no experience: one gets haptic training the other doesn't. Which is very similar to Aidens. And there's also a third group with actual experience in palpating horses. The idea being, the first 2 groups, the group with haptic training would do better than the group with no training… or the group with traditional training. The idea we had for having experts in would be get them to come in and try the simulator, and also have them try the specimen ovaries, and try and pick up some sort of correlation between performance on the specimens and performance on the simulator. Because one of the things that is really hard to do is validate these things. I think we're the only ones working at this in vet medicine. There's lots of people who are working on this in medicine, and at the end of this, they say well there's no ethical way they can test validate this. They can't try untrained patients on… untrained surgeons on patients. So quite a common way is to look for first of all perform when using the simulator, and secondly, looking at the performance of experts and novices, and trying to pick out differences, and see if you can identify an expert's results from a novice's.

S: Are you definitely going to do it with… because you… are you definitely going to do the experiment with the students with horse ovaries or are you doing it with cow ovaries?

A: Its something I was reasonably certain with up until reasonably recently when Stuart was saying that the models we've got are more a sort of amalgamation of both… so I'm not sure. We're going to do it with whatever specimens we can get from a start.

S: what with the foot and mouth thing. But the students will have overall better experience with cows than they will have with horses. And you will be able to get… I mean if you want a group of students… it's a bit artificial, but if you get fourth year students, they'll have less experience than final year students and then you've got Kathryn and Suzie and people. You've got… they're expert, or quasi-expert. So my guess is that cattle would be better. Stuart gave me a copy of Aidens report and I kind of skim read it. I didn't entirely understand what he found

A: I think in the end, he ended up with a lot of numbers, and he didn't really know what to do with them.

S: To be fair it was all rather kind of rushed. To be very meaningful, he really would have… I mean I can't remember how many students he actually had, but maybe a dozen or so in each group. And in reality, my guess is that you need a hell of a lot more. In a way, it was a relatively big effort to get even that number done.

A: I mean the holy grail of this is I suppose… Stuart again was saying it might be possible to get a longitudinal study. That is maybe pay students £50 if and only if they do 5 sittings and look for improvements in performance first of all on identifying follicles on the simulator, and seeing if that again correlates to specimens – horses or cows, whatever's available. So that might be something I'd hope to run next year, but I don't know about availability of students.

S: Well perhaps next year, because we've got the students coming here for a little bit bigger time… even the fact that the fact that they're coming… I mean the detail at the moment doesn't matter, but if we've got them coming for eight weeks in total, but we've got them coming for two week blocks. They start in July August September will be their first block, and then they'll come back kind of January through till May or whatever. So by definition, their general clinical experience will be improving by then, and it would be kind of longitudinal. Basically, I wouldn't give them any money. They'll do what they're bloody well told. But you know all be it that they are somewhat non random group and but I'd be very surprised if the students that Stuart has lined up when they come back from Christmas will be a random group. In fact I can tell you right now it'll be Laura… basically all the blonde glamorous ones… and they're not random, and they would be better than average at this.

A: I think he signed up 20 students, the idea being that 10 would be in the group with haptic training, and 10 would be in the specimen group… well, maybe specimen and haptic just to give them a go on the haptic thing.

S: I think that's the other thing. They'll all be convinced that they've missed out.

A: That's hopefully going to be lined up for after Easter we'll need to wait and see how it goes. What we're also going to do is maybe record movements, and maybe take the data along to someone else to see how it could be used. Maybe as an assessment tool, or maybe as a teaching tool… sort of dragging people along in different directions.

S: Funnily enough, that is you know… its quite interesting now that you actually say that. You've obviously thought that in some ways a lot more than me. One of the most frustrating things about trying to teach someone to do rectal palpation in general… ovaries is an example, but in any other aspect of rectal examinations like intestines and so on. By definition, they put their hand in, and your trying to talk them through it. And when you say, can you feel this, and half the time they just say yes because they're terrified to say no to me apparently… so they just sort of say it, and you don't know what they're feeling. So you don't know whether they actually got it right or wrong, so you never do know for yourself. You say can you feel the nephrosplenic ligament, and they say yes and I'm sure they're not feeling it half the time.

A: There's nothing worse than. I've been doing tutorials this year and my class this year, if you say do you understand that then, they yes, and then call you back 2 minutes later saying well I didn't really get it.

S: To some extent in a way you would be able to… if you were able to… almost looking at the other side of the screen as it were. If you had a screen that they couldn't see then you would know. Not so that you could chastise them, but… and if it could be interactive if you could say, no no that's not it and literally move they're hand and drag them to the point.

A: Funny you should say that because we again talked about this thing sort of again in 200 year time or whenever… unless the cost of these things comes down then it will be. But the sort of thing where you might have a tutorial room with all these devices in it. And you might have a tutor that can observe the people doing the different examinations, or it might be actual surgery… and actually choose to interfere with them. So say no you're doing this wrong. What about over here? Go over this way. And again that's something we have a very basic model of working for the whole algorithms necessary to drag people through a series of steps are more complicated than I initially thought.

S: Well I can see that… that really would be fantastic, because its kind of like the unsatisfactory element, because you just don't know if they are getting it or not.

A: That's actually quite interesting, because it gives me an excuse to go back to that. I quite enjoyed that. It's something that no one else has done so far, so it's quite good for my PhD.

S: The other kind of thing is when your trying to teach them to… it's a different kind of thing all together really… but when your trying to get them to be able to find a heart murmur. A horse's heart is quite a difficult thing to osciltate. I.e. using a stethoscope. So you kind of say put your stethoscope on here and then say can you here it, and they say yes. And they can't here it at all. In other words, you can never check if they're actually doing it. It's a slightly different thing although in a way its not actually entirely different because their ability to hear it or not is to be able to recognise the noise. Actually there are some crummy recordings of the noise. When you here the whoosh, that's the murmur and there are some people who would be good a going lub-whoosh-dub lub-whoosh-dub, or lub dub whoosh depending on where the murmur comes. But being able to detect it a lot of the time is being able to put their stethoscope on the right part of the horse's heart. The horse's heart is a big thing and it's elbow is in the way and things like that you know where your heart.

A: We felt a vein here (horses chin) we got to try it on I think it was a pony.

S: You mean when you were feeling its pulse… yes.

A: yes

S: That's a classic… Funnily enough, and I'm thinking of third year nightmare classes that we have. The basic physical examination of a horse. And I use to have these crazy objectives on these things, and now if they're able to count the pulse, I consider that to be a major triumph. But it's amazing because you can… I mean the thing is it's slightly thicker than that - and a horse's pulse is quite slow – which puts them off because they're use to their own pulse or maybe they aren't… It's quite slow so that puts them off. You literally put their finger on it and say can you feel it now, and they say no. You want to say… well leave the course. You think how can't they feel it. It's under their fingers. I don't know how you could produce, I mean with pulsing…

A: We have… well a slightly scary thing that happened I suppose. We had a guy… he seems to know everyone, so you might know him… Martin Bartos? I think he knows Stuart. Well he's working with CMCSM… I think it's CMCSM. It stands for something strange like algorithms, computing science and maths and medical science or something. But they're working between Strathclyde and Glasgow and they're basically interested in interesting and cost effective ways of improving the health service. They work very much with human medicine I think, but I've had Martin in a few times to try things out. He's great for ideas. He's always got good ideas about needlework and what have, and he's always free to come in to try things. Well one of the things he immediately said when he tried the simulator was… oh, I can feel a thrill there, and we didn't expect that. So I suppose one of these things we've really got to be careful when you try and model something is introducing artefacts that you don't really know are there… so we can definitely do thrills. Whether we can get them without thrills is another matter.

S: I mean think there's lots of… in a way, quite a lot of things that are physical feeling like emphysema or pitting edima or… these are things that you show to third year students and say if you press it, it leaves a big pitting… what would be really neat would be when eventually you have a ****. I mean I remember going to Michigan state vet school, and every single seat in the place has got a PC at it. And the lecturer says Andrew you show what you consider to be the aorta, and you have to show him the aorta….. if you took it a stage further with physical things, or like precordial thrills or fremitus. I don't know if that's a word you've heard. Diagnosing a stage of pregnancy in cattle.

A: How do you spell that?

S: F-r-e-m-i-t-u-s. Emphysema is gas under the skin, so it feels like a bubble pack. Ideally, every student should have felt emphysema. It's not pulmonary emphysema. You here old men saying… it's a sort of old fashioned thing. But subcutaneous emphysema, and you know, having as physical module where all these things are… because they all just are physical experiences… or skytees, which is like when you've got fluid in your belly. Its just movement of fluid in your belly, but until you've actually felt it a few times…

A: One of the things that has come up a few times is access to the material so if it's a rare case, then how are you going to feel it regularly before you go out and practice.

S: Yes.

A: I think Martin was saying that he always use to get – when he was in wards – students going round to say go up to floor whatever… there's an interesting case of whatever… so go and feel it, or go and hear it. SO you always use tell each other the interesting cases.

S: I'm sure it's not good for the psychology of the patient. Do you mind if I poke at your swelling… so there would appear to be a lot of scope… and it's not just limited to vets. You've got vet nurses, and things like that too. You know its kind of biggish.

A: So again I think we're really the only people anywhere who are doing haptics with vets. There's a growing amount of medical haptic stuff, but its all medicine and they all seem to be based around the same procedures. Usually MIS stuff, and there's a few surgery applications that I've been quite impressed with, but it's really just us and Rutgers University who are doing palpation. And I suppose they did something quite similar. They did prostate in men. Palpating prostate in men for much the same reasons that we're doing our stuff here. Also, breast cancer palpation, and liver palpation or something.

S: Well in an animal – it's quite funny – 20 years on, we're diagnosing pregnancy in a dog uterus by feeling the puppies. But 3 or four weeks pregnant, we now put ultrasound on it. But a lot of the traditional palpation methods are dying out. Which is a shame really but they're… some of them at least that are less important, but some of them will always be important. Detecting these in physical abnormalities in horses with acute abdominal crises and colic, well the thing that would be good for that would be A) it would be a really really really useful thing to produce, and B) it would be relatively easy to get funds for because there's a big welfare issue. Horse in agony.

A: I'll tell that to Steve…

S: It might well be a 3 year project for someone… I don't know how difficult it would be. And I think again because all the hardest part… or some aspects of the hardest part have been done by yourself, then there would be… your only looking at maybe 5 different kind of overall scenarios. I mean in the ideal world, you have the whole body organs, there, and you have the bits that are normal beside the bits that are not normal. And you'd just put your hand in and say that is the large intestine and it's filled with ingesta. That would be fantastic.

A: As I say, Steve's always looking for new ideas. We are looking to get more people in. Just now the problem's been we have 2 devices and we have four, or five people working on it. So we are quite tight for equipment. But 2 of these people are just about at the stage of writing up, and we are looking for new projects and ideas. We've talked to the graphics lab who are doing various stuff with visible human stuff, and lots of 3d scanning, and lots of interesting stuff like reconstructing a face from a skull, and all sort of things like that. If you're going to build up a face from muscle, and bone, it would be really nice to have a haptic interface to that as well. So we are looking for new ideas, and any ideas you do have, we're always interested.

S: Sure.

A: So I think that's me. Thanks for that.

S: I won't psycho-analyse why you used a red pen…

Interview 2 Transcript

The interviewer (Andrew Crossan) is A. All sections that were unclear from the tape are marked with the symbols ****.

A: It's really just to find out about the course, and what's being taught, and how a Virtual Reality simulator would fit into ... well any course in general, but specifically an ovary palpation training course. Just a few question about the course, and how it's taught, and a few questions about requirements at the end. So firstly for the vet course, what are the course requirements for first through to final year for ovary palpation.

K: First years get gross and microscopic anatomy, so they get like a tract and stuff like that... first and 2nd year.

A: An anatomy, lab?

K: Yes, its across the road - I think you had a look at some stuff on trays didn't you.

A: I think it was cow tracts.

K: Yes they get that for all species in first and second year… and they get histology and... but that's only really being given specimens. I don't really...

A: Right.

K: They would just be to study the anatomy and the like.

A: What form do the labs take? Do they have a specific task to do, or are they just presented with the tracts, and told this is...

K: A lot of the thing with tracts... in the dog stuff and the cat stuff they would dissect them.

A: I was in a dog lab, but I was trying to face the other way.

K: Obviously they can't do that with horses or cows...they just get given the tracts and the trays, and whatever.

A: Is it very much 'Here's your tracts, and...

K: Yes - you get to feel it, and poke around, and you have a go at dissection.

A: So…

K: They don't really get a lot more until final year. Then in final year, they'll get some classes there. But during the 5 years of the course, there's a component of seeing practice called **** , so during that time, people will have been expected to go out with vets, and have palpated...

A: So outwith the vet school.

K: And in they're final year, they'll get rectalling classes.

A: Ok What about theory wise from the point of view of lectures, or...

K: They'd get anatomy and physiology lectures in first and second year.

A: I'm just writing this down because it didn't work the last time.

K And then they get separate reproduction stuff in fourth year. So they a course about the actual hormonal events that... so control of reproduction if you like. That sort of stuff.

A: My next question was really how the course objectives are achieved, and that's pretty much...

K: Yes - the students do get a list of aims and objectives for the course at the start of the year.

A: Would it be possible to get copy of that... not necessarily just now.

K: Yes - you'd probably have to speak to the relevant departments of first year and second year.

A: Are they on the web.

K: I don't know. And they get an aims and objectives booklet that gives them and idea of what they're supposed to know.

A: I've got a blue booklet that I've still to get all the way through... So what would you say the core skills are in ovary palpation. So what skills would they have to learn to perform the examination.

K: In a real animal?

A: Yes.

K: Well I guess you'd need to know the actual anatomy. Where things are relative to the ovary. A lot of the problem is actually trying to find the things...

A: That's something that Monika brought up...

K: So it's trying to find the ovaries relative to where the uterus is... because your kind of main landmark would be the cervix. You can find the cervix pretty easily in most cases and then you try and work things out from there. So follow it forward to the end of the uterus, and then try and work out where your ovaries are in relation to the uterus. So its kind of knowing where things should be so you can find them I guess.... so manual dexterity I guess would be a core skill. If that's the sort of thing... just actually the ability to...

A: The kind of motor skills?

K: Yes.

A: So what about when you've actually found them, and you're holding them.

K: So when really just working out what's happening.

A: Surface features?

K: Yes. It would just be just actually touching it... it would be the actual size of it and kind of consistency of it, and then the sort of surface struct... the surface texture almost. I mean is it a kind of smooth bean shaped thing, or has it got something sticking off of it, and if it has, work out what it is. Like follicles are always kind of smooth as opposed to CLs which are very fleshy feel.

A: So another thing that Monika said was a lot of the work is now done with ultrasound.

K: yes.

A: What would you say the hardest thing to teach a student is.

K: Finding the ovaries.

A: Say you take a student along to a few labs, how fast do hey pick that sort of thing up.

K: Relatively quickly I suppose. But its just the kind of practice aspect... you know, the more you do the... Its very difficult when the students got they're hand in the cow to try and then... obviously you don't know yourself where it is, so your trying to explain where its most likely to be.

A: Right. It’s more an experience thing?

K: Its so the students are happy that they can palpate the uterus, and then work out from that where...

A: Right so just location of ovaries... and its just kind of experience in working out where things are.

K: Yes.

A: Again, you've probably answered this to some extent earlier, but what part does external learning... learning outside the uni have in the course. Are the students expected to go out and get training in the field.

K: Yes, they have to do a certain amount of practice each year, and particularly in fourth and final year... I can't remember whether its 26(???) weeks during the holiday time of EMS which is spending time with vets, and hopefully getting to do that sort of thing.

A: Do they get to choose what they specialise in?

K: Sort of a fairly balanced amount of it. Obviously they'll have a particular interest and they can do more of that but they have to do a sort of core amount.

A: Right. And how much... are students expected to go to the library as well, and read up on...

K: Yes.

A: Could they pass the course if they didn't go to the library, or is it just...

K: They get printed notes for a lot of what they do, so they get provided with handouts. And the practical s reinforce this and should be enough to get by, but its the sort of reinforcement.... ****

A: So they would be expected.

K: They'd be expected to use the library yes. But its also on the computer cluster, there' a computer aided learning program. There's different programs for all the different courses there.

A: We've got a contact in computing science who knows a bit about that, so I'll have a look at that. So how is the time divided into for ovary palpation training from first through to final year?

K: First year it’s a very small amount of anatomy. Literally just 1 or 2 sessions of here are your tracts. And they don't get anything in 3rd year and in 4th year, it would be lecture based.

A: In a course on reproduction or something?

K: Yes. Which is... I don't know how many hours of lecture time. Maybe 6 or 7 hours of lectures.... something along those lines. They'll be on a ll aspects of reproduction.

A: Presumably on different animals.

K: Yes.

A: Is there quite a bit of overlap between the animals?

K: Yes. Once you've found out stuff about the cow ovary, you'll be expected to apply that to sheep. Obviously you don't palpate sheep ovaries, but you know the kind of theory. Then in final year, there's quite a bit in final year. You get 2 sorry 3 routine farm visits to do rectalling. Then you get one session purely on scanning. Then you get a tutorial which is based on tracts and scanning tracts.

A: So what would you like to see added to the course. Is there something you'd like to see done that you....

K: More practical. More practical sessions. With animals rather than tracts.

A: In vivo?

K: Yes.

A: So how would you fit that in? Would that be an extra part of the course, or would you cut something out of the course. Is there space to incorporate it into the course just now?

K: You could probably get more in final year. You could probably maybe cut out some of the practical classes we do at the moment. The final years get a lot of tutorials and presentations and stuff, and you could probably instead of one of those do another session. It would probably be easier to get it into final year, but it would be more difficult earlier in the course because of the number of people on the course. Like in final year, you get small groups of people going through a rotation at a time. So you maybe get 6 or 12 people at a time in final year. Where as earlier in the course, your dealing with all the students together, so you maybe get 80 or 100 students in one go which obviously is difficult to co-ordinate. The 4th year course is very very busy anyway. They have a lot of stuff to do in 3rd year

A: Do they all generally do the same modules.

K: Yes. It's actually changing next year. In 4th year, they'll still do the same course, but in final year, they'll get to choose between equine or farm animal. They'll still do a basic core course on equine and farm, but if they want to do more farm, they can choose to.

A: Right, I've seen vet medicine split into large animal and small animal, but is there also farm animal as well as that.

K: Well, hey kind of split it up into small animal, equine and farm animal. In final year, they get to choose on whether they do farm animal or equine as a... I don't know what the word is. They'll all get a basic core course on it, but then they'll have to choose. Half the year will do farm, and half the year will do equine as a kind of extended course. That starts next year.

A: OK. Again, is there anything you think could be improved on.

K: As a whole in the course?

A: Yes.

K: More practical, and just really a greater opportunity to do more scanning, and more rectalling with people there to help you. The difficulty you see in practice as a student is your going out with a vet, and the vets doing a visit for a farmer. It's maybe a timed visit. The farmers paying by the hour, and if he's taking ages because he's trying to explain something to a student, the farmer's obviously going to get a bit annoyed. Its difficult sometimes when you see practice during visit, that although you get to see the visits, you maybe don't get the time spent with you that you do here. So if there was a greater opportunity either spends more time with either vets in practice, and they could kind of co-ordinate so that the vets charge differently for those students. If there was some way you could liase differently with practitioners that would give students more opportunity. Or if students here had more chances to go up to the farm, and get more classes.

A: If you had your ideal simulator, what requirements would it have?

K: It'd have to... I think you'd have to have some sort of cut off mechanism when the students were exerting too much force.

A: So some method of warning the user...

K: Because some animals, particularly horses when you rectal then can get quite sore. SO some way of when the force is too much... a kind of cut off point.

A: SO some that's something we've kind of being working on.

K: And the same sort of thing with duration. There's a tendency with some students, we take them out to the farm, and they can't find the ovaries, and they spend hours and hours and hours trying to find ovaries, and they don't really realise that they've maybe had they're hand in the cow for 10 or 15 minutes. Obviously that's too long, so if there's some kind of time limit you could build into it so that after that you could say if you've not found them it doesn't matter. Keep practicing until you can do.

A: So as well as the actual skills, you've got the pre-procedure and post procedure constraint as well.

K: As well as actual do what you have to do , and find out what you have to find out, you've also got to be wary that you're not going to hurt the cow, or if you've got an early pregnancy your going to palpate, there's obviously a risk of abortion, so having some sort of cut off like duration or force. I think as well if you could find some way of mimicking a peristatlisis you get in a normal cow. SO when you were doing it, if you remember the kind of gut movement you got on your arm. f there was some way of mimicking that. Because if the cow was sort of clamped down on kind of rectum around your hand because you're hand's squashed down like this trying to do anything.

A: So the environment round, not just the ovaries.

K: Yes I think that there's lots of different variations that you could have. So you could have... because there's so much variation between cows. Like a heifer that's never had any calves, is going to have a really tiny tract, and really tiny ovaries that are really small. The whole thing fits in your hand. Where as if you've got a suckler cow that's maybe had 10 calves, her tract is going to be bigger anyway, and its a much bigger cow and you've got to stick you're hand in a lot further and grope around a lot more to find it. So if there was some way you could adjust the program...

A: SO adjustable size and position.

K: Yes, as well as what's on the ovary.

A: How would it fit into the course. I've heard - can't remember who told me about it - about a black box simulator. I don't know if it use to be used or is still used or if its still part of the course. So how would a simulator fit into the course?

K: I think it would be use in 1st and 2nd year with the anatomy classes to do it then. But it would also be very useful in 4th and final year. When they do they're clinical stuff. Particularly in final year.

A: Would there be time in the course.

K: Yes.

A: Last question. Is there anything else that I've missed.

K: No... not really. I think it is just really trying to emphasise that it really is like riding a bike. its just really practicing and practicing and practicing and practicing, and doing hundreds and hundreds of cows before you can feel happy enough in your skill of doing it. It's trying to expose students to as many different cows or tracts as you can. Which obviously is difficult because we don't have the number of cows here to do that. That's the main constraint we have a t the moment. There's only 80 cows at the farm, and you can only rectal them 3 times a day...

A: What about horse wise. Does the vet school have horses?

K: They have, but again that's really limited, because horse have much more **** rectum anyway, so your on to a non starter right from the beginning, and there's only a very limited number of horses. Maybe 6 or so holding ponies that people can get practice with. But needless to say, they can only get rectalled a certain number of times. It's difficult especially with horses. They get far less horse exposure than they do with cows. I didn't rectal a horse at all until my finals, which was nice. A great way to start.

A: Well I think that's just about it. So I'll say thanks a lot.

Interview 3 Notes

- What are the course objectives for teaching equine ovary examination working through from 1st to final year?

Provide general background and theory to the students. Not necessarily train them to be proficient ovary palpators.

- How is each of the course objectives met?

2nd year, students are provided with some theory ~ 2 lectures and also, 1 anatomy lab (in vitro training) 3-4 hours. Students can specialise in their final year, and some may perform a few actual (in vivo) ovary palpations. When newly qualified, vets learn in their first few weeks of actual practice. Different universities provide different training. For example, Edinburgh provides in vivo training to students.

- How well can each of the course objectives be met?

General theory is provided – the course does not try to train students as a competent ovary palpator.

- What are the core skills involved in the procedure?

Locating the ovaries is important. Distinguishing between objects (eg intestine, uterus, ovary). This is done by size and consistency of the object. Once found, no core skills - just palpate. Surface properties for the ovary:

Corpus luteum – Thick walled, ridged

Follicle – Thin walled, soft

Ovary – Hard fibrous

- What is the hardest thing for the students to learn?

Locating and identifying the ovaries. Palpation learned through lab and practical experience. Examinations are mostly ultrasound, although some palpation skills are still needed. Features on a horse ovary are often partially submerged, which means that ultrasound is better at detecting these features. Cow ovaries are more often palpated.

- Is training commonly provided from outside university sources?

Yes – Interested students can try and get experience at a vet practice. Again training can be provided in the first few weeks of a job.

- What part does self-learning through (library work etc.) play in training?

Only so much theory can be learned through library work. Students might talk to an experienced horse vets about their experiences.

- How is time divided into lectures/tutorials/lab work/practical experience for each year?

2nd year – 2 lectures, 1 (3-4) hour lab. A 'black box' simulator is not in general use. Look away and feel tract is currently the best method.

- What other areas would you like to see included in the equine ovary palpation course objectives?

More opportunities for in vivo training for the students.

- What training could be improved upon?

More in vivo training. Vets who choose to specialise in equine vet medicine may get an afternoon training. Quantity of labs is important. Maybe more shorter labs in order to fit into the timetable.

- What would be your requirements for a simulation to be incorporated into the course?

Adjustable. The simulator should be flexible. For example adjust ovary size, location, surface properties, stage of cycle, model tumours, change the stage of pregnancy. The simulator must include the environment. One of the main problems of in vitro training is the tract is removed from the environment.

- How do you feel a simulator could fit into the ovary palpation training course?

Time could easily be provided for students who choose to specialise. Maybe remove a sheep lab, and insert equine ovary palpation lab.

- Can you think of anything else to discuss?

We're maybe ignoring other palpation areas. Colic is a big problem in horses, urinary tract palpation. Needle simulation would not be as useful. A student can practice on an anaesthetised animal. A needle simulation providing different skin types might be useful. For example Pig – rough, cat – soft, horse – soft, cow – tough. Palpation for a vein is another useful skill. A MIS simulator might also be useful.

A.3 Classification of Data from the Interviews

Category CI - Reproductive Course Information

CI 1: M: Course Objectives are to provide general background and theory to the students. Not necessarily train them to be proficient ovary palpators.

CI 2: M: General theory is provided – the course does not try to train students as a competent ovary palpator.

CI 3: M: 2nd year, students are provided with some theory ~ 2 lectures and also, 1 anatomy lab (in vitro training) 3-4 hours. Students can specialise in their final year, and some may perform a few actual (in vivo) ovary palpations. When newly qualified, vets learn in their first few weeks of actual practice. Different universities provide different training. For example, Edinburgh provides in vivo training to students.

CI 4: M: Palpation learned through lab and practical experience.

CI 5: M: Interested students can try and get experience at a vet practice.

CI 6: M: Training can be provided in the first few weeks of a job.

CI 7: M: Only so much theory can be learned through library work.

CI 8: M: Students might talk to an experienced horse vets about their experiences.

CI 9: M: 2nd year – 2 lectures, 1 (3-4) hour lab.

CI 10: M: A 'black box' simulator exists but is not in general use.

CI 11: M: Look away and feel tract is currently the best method of ovary palpation training.

CI 12: M: Vets who choose to specialise in equine vet medicine may get an afternoon training.

CI 13: M: One of the main problems of in vitro training is the tract is removed from the environment.

CI 14: L: different vet schools have different approaches.

CI 15: L: there is a kind of ethical issue that does arise. Is it fair to abuse the pony's backside basically and all in the interests of vet students ?

CI 16: L: currently, all the students are expected exactly the same course, although all that's going to change.

CI 17: L: first and second year subjects that would relate to this subject are specifically anatomy of equine reproductive tract, male and female, and physiology and it would be true to say that the physiology is largely taught as a kind of generic thing.

CI 18: L: there are fundamental differences (between species) you can compare and contrast, but it would certainly be at a pretty superficial level.

CI 19: L: the objectives would be more you know to know the overall hormonal events that control and fundamental things about anatomy, as a basis for when it comes to do ovarian palpation.

CI 20: L: knowing structures, knowing functions really, but not at any great detailed level.

CI 21: L: the methods of manipulating the reproductive cycle i.e. with various administered medications

CI 22: L: The only third year subject that would even vaguely encroach.

CI 23: L: essentially when we would get into a little bit more depth and currently there are 8 soon to be reduced to 4 lectures on equine reproduction. So that includes male and female, mostly female.

CI 24: L: so that would then include a kind of hour of kind of normal cycles and manipulation of cycles and then an hour of kind of fertility practices, you know ways of maximising things, and ways of manipulating ovulation.

CI 25: L: Using ultrasound

CI 26: L: the objectives of fourth year are really to make people appreciate the main focus of fertility management in the mare because bottom line is that's what reproduction in the mare comes down to.

CI 27: L: What it comes down to is 2 almost discreet areas of it, or maybe 3. Manipulation of the ovarian cycle, and prediction of ovulation to optimise the time of insemination - so mating, that's what we're just been talking about. Number 2, early pregnancy diagnosis. Which again tends to be done by ultrasound, very early - 15... 16... 19 days. And then the third main component of equine reproduction is investigating reasons of infertility, in other words, why is it not getting pregnant. And those are far and away… those would be the 3 main components, and that's probably what the objective of the fourth year course is.

CI 28: L: So basically, they learn about these 3 things, and then there's a sort of catalogue of... currently there's no hands on skills.

CI 29: L: The hardest thing to teach probably is early pregnancy diagnosis… so that's a uterine event obviously.

CI 30: L: That's probably harder to teach than has this horse got developing follicles. Is the horse near ovulation . Has the horse ovulated.

CI 31: L: There are then a short catalogue of specific diseases and things that the students would be aware of.

CI 32: L: Well the point is to illustrate a concept. Literally to illustrate physiological events to make them graphic, but not to teach them skills.

CI 33: L: Currently in the final year then is get all the students experience of palpating ovaries and uterus in mares. Realistically, I think we nearly achieved that.

CI 34: L: Most students will have had they're hand in several mares and that would be to try to get them to learn some physical skills.

CI 35: L: As it so happens, this next year, this is the first year we're going to be having species tracking so not all students will come onto the equine course. So we'll have a smaller group of them and theoretically at least, they will be the ones that will really want to do horses.

CI 36: L: They're (The students numbers) growing but now capped. I mean basically we gone in a sort of… well certainly in a 10 year thing from having 60 to having 100.

CI 37: L: overall I suppose what I would say about in terms of an objective of the teaching at undergraduate level, the objective or expectations of their skill is pretty limited.

CI 38: L: I wouldn't really expect the students to be anything like as competent in this specific subject as I would expect them to be in lameness examination, because even within the very specific sort of horse vets, huge numbers of them do not do this.

CI 39: L: if they're going to become really skilled in it then it's a postgraduate subject for them.

CI 40: L: At the moment it would be true to say that first second and third year are all lecture based, and nothing else.

CI 41: L: Fourth is lecture based.

CI 42: L: first through to fourth year is lecture based purely.

CI 43:L: nearly all vet student lectures will be reinforced with visual materials i.e. slides of specimens or ultrasound videos and that sort of stuff.

CI 44: L: in fifth year final year there are practicals i.e. a horse back side and a demonstrator reinforced by simultaneous ultrasound and palpation.

CI 45: L: And we have - which your welcome to take away if you want – a couple of CD ROMS on equine reproduction and a video on early pregnancy diagnosis.

CI 46: L: but my view is the objectives are of an awareness at a bit more than a conceptual level, but not massively more, which I think we achieve.

CI 47: L: we don't set ourselves very ambitious objectives probably so we do achieve them.

CI 48: L: the biggest limitation is probably the group numbers and as I said we've actually got something in hand to try and effectively try and reduce the number of students that want to do a practical.

CI 49: L: the fourth year lectures are provided by an outside source… i.e. I guy who makes his living doing this and nothing else. That is a guy who makes his living doing equine reproduction work and nothing else. And he comes on a hired gun basis and gives the student lectures. And he's highly experienced.

CI 50: L: In terms of the handful of students going on placements, and it is literally that… maybe 5% of students who really want to get into equine reproduction in a major way.

CI 51: L: We would get them (interested students) placements in practices where this is a major part of the practice activity.

CI 52: L: essentially probably self-learning would probably be the relatively few who are interested would make use of the CD ROM type stuff to had. And to be honest, its commercially produced and purchased, and its not of sensational quality.

CI 53: L: Essentially all it (the CDs) is, is digitisation of a course which is run for postgraduate level for vets in practice, and its basically digitisation of some lectures, some of which are very boring. Many of which are very boring.

CI 54: L: Well overall, it doesn't comprise a very large component of the equine teaching.

CI 55: L: the student hours on the subject year 1 through to 5 would probably be about… 30 now that is out of something in the order of believe it or not 2800 hours.

CI 56: L: I suspect its probably similar to Bristol, but probably less than some of the other places… it might be at the maximum it would be twice that so its still pretty small… Less than 2%.

CI 57: L: there are 25 to 40 clinical disciplines.

CI 58: L: In equine lameness for example, the hours would be in that probably – could be 100… maybe 120. But that would be about right because its massively more important subject. It's probably relative to its own importance.

CI 59: L: That figure of 30 hours for these students who are… we're calling it sort of equine tracking students, that figure's going to go up for them. So the tracking is only going to apply to final year but I would say that that figure will probably only go to 50 hours. For them, they are going to get massively more.

CI 60: L: Currently each of the final year vet students comes here in equine hospital as part of their final year, which is essentially all… in the final year, there's no lectures, its all clinical hospital training.

CI 61: L: Each student currently comes for 2 weeks. Next year the 40% who have chosen to do large animals… sorry, who have chosen to do horses as opposed to all large animals, are going to do 8 weeks so basically, we're going to have more time to teach them these physical skills.

CI 62: K: First years get gross and microscopic anatomy, so they get like a tract and stuff like that .... first and 2nd year.

CI 63: K: Yes they get that for all species (anatomy lab tracts) in first and second year… and they get histology and... but that's only really being given specimens.

CI 64: K: They would just be to study the anatomy and the like.

CI 65: K: A lot of the thing with tracts (in anatomy labs)... in the dog stuff and the cat stuff they would dissect them.

CI 66: K: Obviously they can't do that with horses or cows...they just get given the tracts and the trays, and whatever.

CI 67: K: you get to feel it, and poke around, and you have a go at dissection.

CI 68: K: They don't really get a lot more until final year. Then in final year, they'll get some classes there. But during the 5 years of the course, there's a component of seeing practice called **** , so during that time, people will have been expected to go out with vets, and have palpated...

CI 69: K: And in they're final year, they'll get rectalling classes.

CI 70: K: They'd get anatomy and physiology lectures in first and second year.

CI 71: K And then they get separate reproduction stuff in fourth year. So they get a course about the actual hormonal events that... so control of reproduction if you like. That sort of stuff.

CI 72: K: the students do get a list of aims and objectives for the course at the start of the year.

CI 73: K: they get an aims and objectives booklet that gives them and idea of what they're supposed to know.

CI 74: K: they have to do a certain amount of practice each year, and particularly in fourth and final year... I can't remember whether its 26(???) weeks during the holiday time of ems which is spending time with vets, and hopefully getting to do that sort of thing.

CI 75: K: Sort of a fairly balanced amount of it. Obviously they'll have a particular interest and they can do more of that but they have to do a sort of core amount.

CI 76: K: They get printed notes for a lot of what they do, so they get provided with handouts. And the practicals reinforce this and should be enough to get by

CI 77: K: They'd be expected to use the library yes. But its also on the computer cluster, there' a computer aided learning program. There's different programs for all the different courses there.

CI 78: K: First year it's a very small amount of anatomy. Literally just 1 or 2 sessions of here are your tracts. And they don't get anything in 3rd year and in 4th year, it would be lecture based.

CI 79: K: Maybe 6 or 7 hours of lectures (in 4th year reproduction course).... something along those lines. They'll be on all aspects of reproduction.

CI 80: K: Once you've found out stuff about the cow ovary, you'll be expected to apply that to sheep. Obviously you don't palpate sheep ovaries, but you know the kind of theory.

CI 81: K: Then in final year, there's quite a bit in final year. You get 2 sorry 3 routine farm visits to do rectalling. Then you get one session purely on scanning. Then you get a tutorial which is based on tracts and scanning tracts.

CI 82: K: in final year, you get small groups of people going through a rotation at a time. So you maybe get 6 or 12 people at a time in final year. Where as earlier in the course, your dealing with all the students together, so you maybe get 80 or 100 students in one go which obviously is difficult to co-ordinate (for practical work).

CI 83: K: Its actually changing next year. In 4th year, they'll still do the same course, but in final year, they'll get to choose between equine or farm animal. They'll still do a basic core course on equine and farm, but if they want to do more farm, they can choose to.

CI 84: K: Well, they kind of split it up into small animal, equine and farm animal. In final year, they get to choose on whether they do farm animal or equine as a... I don't know what the word is. They'll all get a basic core course on it, but then they'll have to choose. Half the year will do farm, and half the year will do equine as a kind of extended course.

CI 85: K: It's trying to expose students to as many different cows or tracts as you can.

Category PCD - Perceived Course Deficiencies

CD 1: M: More opportunities for in vivo training for the students.

CD 2: M: More in vivo training.

CD 3: M: Quantity of labs is important. Maybe more shorter labs in order to fit into the timetable.

CD 4: L: their (Universities with reasonably sized pony herds) students probably get slightly more experience than our students do.

CD 5: L: it's very costly to do that (Keeping a herd of ponies)

CD 6: L: having a little herd of ponies is almost worse than having none because we've got so many students that there's nothing worse with students than when one gets the chance to do it, but another one doesn't they feel madly cheated.

CD 7: L: and numbers do make these things very much much more difficult to be able to give them all experience.

CD 8: L: And because the horses that come here – I don't know if Antony showed you round its not quite finished yet - they all come in and they're owned by people so you can't really turn a student loose on them. So you can't give the students any sort of experience.

CD 9: L: getting their level of skills at basic pre-breeding assessment and or fertility assessment of mare so yes we would for a certain proportion of students, getting them more competent in those… what are everyday procedures if you take this up as a career.

CD 10: L: Artificial insemination is becoming a lot more common in horses than it was.

CD 11: L: at the point of graduation, a technician who would have no qualifications essentially at all is likely to be better than a vet (at artificial insemination).

CD 12: L: So that is the main way we can improve it. By providing time and facilities. Now we can't provide 100 horses, or 50 horses, or 40, so having teaching aids would be useful.

CD 13: L: Follicular development and pregnancy diagnosis… particularly early pregnancy diagnosis could be incorporated into that (simulator) and that would be second year type subjects.

CD 14: K: More practical (should be added to the course). More practical sessions. With animals rather than tracts.

CD 15: K: You could probably get more (in vivo practical classes) in final year. You could probably maybe cut out some of the practical classes we do at the moment. The final years get a lot of tutorials and presentations and stuff, and you could probably instead of one of those do another session. It would probably be easier to get it into final year, but it would be more difficult earlier in the course because of the number of people on the course.

CD 16: K: More practical (should be added to the course), and just really a greater opportunity to do more scanning, and more rectalling with people there to help you.

CD 17: K: The difficulty you see in practice as a students is your going out with a vet, and the vets doing a visit for a farmer. It's maybe a timed visit. The farmers paying by the hour, and if he's taking ages because he's trying to explain something to a student, the farmer's obviously going to get a bit annoyed.

CD 18: K: Its difficult sometimes when you see practice during visit, that although you get to see the visits, you maybe don't get the time spent with you that you do here. So if there was a greater opportunity either spends more time with either vets in practice, and they could kind of co-ordinate so that the vets charge differently for those students. If there was some way you could liase differently with practitioners that would give students more opportunity. Or if students here had more chances to go up to the farm, and get more classes.

CD 19: K: Yes I think that there's lots of different variations that you could have. So you could have... because there's so much variation between cows.

CD 20: K: (exposing students to cows or tracts) is difficult because we don't have the number of cows here to do that. That's the main constraint we have at the moment. There's only 80 cows at the farm, and you can only rectal them 3 times a day...

CD 21: K: They have, but again that's really limited, because horse have much more **** rectum anyway, so your on to a non starter right from the beginning, and there's only a very limited number of horses. Maybe 6 or so holding ponies that people can get practice with. But needless to say, they can only get rectalled a certain number of times. It's difficult especially with horses. They get far less horse exposure than they do with cows. I didn't rectal a horse at all until my finals.

Category SI - Simulator Information

SI 1: M: Adjustable. The simulator should be flexible. For example adjust ovary size, location, surface properties, stage of cycle, model tumours, change the stage of pregnancy.

SI 2: M: The simulator must include the environment.

SI 3: M: Time could easily be provided for simulator work for students who choose to specialise. Maybe remove a sheep lab, and insert equine ovary palpation lab.

SI 4: L: Probably to use examples of extremes. Here is a non-cycling mare with no ovarian structures. Here's your mare on the point of ovulation. But you would be needing to teach all of the students... or I don't think it would be appropriate that 100 students would need to go and know those skills in a way they would have to go and apply them.

SI 5: L: the absolute recreation of real life sensations in the model isn't massively important. We're not going to try to train out and out – or I wouldn't but you might – I wouldn't have thought that your trying to think that this is absolutely real life you've recreated, and that really exactly how it is.

SI 6: L: Its kind of like this is sort of what it will be like, and this is scenario A, B, C, D, E, and F, and if they do it often enough, they will get the general feel.

SI 7: L: I think if it (the simulator) was to be absolutely ideal, I suppose you would have every scenario of ovarian physical events that you could have, so that matter ovarian and uterine.

SI 8: L: When I say every scenario, there are not that many, and they may well all be on there. But you would have everything from small hard ovaries with no follicles through to large twin ovulations, and then probably the odd scenario such as the relatively rare disease states of the ovary such as ovarian tumours. And certainly, would as ideal have it sort of various stages of pregnancy, particular early pregnancy with appropriate ovarian events that go along with that. So in other words, correlating early pregnancy, and in the same simulation, having ovaries that would match with that pregnancy.

SI 9: L: at the moment its in a vacuum and when you put all the other bits – the abdominal organs in – then I guess if you like that that would make it more real.

SI 10: L: if you wanted to take it (the simulator) to its real extreme… maybe the mare's straining on your arm.

SI 11: L: if you can slide it along it kind of is progressive, and you can feel in your very fingers… kind of if you slide it to 72 hours before ovulation and then as you slide it closer to ovulation, its changing in front of your hand. I'd imagine that would be a very useful thing.

SI 12: L: … presumably, you could set up self-assessment once they've done the training program a few times or whatever. And that from an educational point of view, that would be great.

SI 13: L: but certainly in the simulation thing, what could be there that could be confused with ovaries.

SI 14: L: well the ideal would be that there was ultimately at the top level, if you like something that included these other confounding things. But that might be for final year or even post grad people where as second year students could learn a lot from your slider and ovarian development. So I think things would be slightly different for different groups.

SI 15: L: Well I think you would use it (the simulator)… or could use it probably as a supplement to second year. I.e. ovarian follicular development probably in particular.

SI 16: L: You could use maybe similar materials, but maybe a little bit more advanced and correlate them particularly with ultra-sonography. I think that would be really neat, and if there was a side by side a sort of as you put your hand on this thing that's developing as you slide the scale along the ultrasound appearance that would correlate with that going on side by side… would be neat.

SI 17: L: And in final year, I think you would use it as a learning tool to try and say… I think you probably would create scenarios of problem solving.

SI 18: L: But I think it would be in a kind of problem solving tutorial.

SI 19: L: It seems to me that combining something that's a physical sensation with a visual… to my mind that would be… just intuitively, hat would reach people a lot.

SI 20: L: If you had a screen that they couldn't see then you would know. Not so that you could chastise them, but… and if it could be interactive if you could say, no no that's not it and literally move they're hand and drag them to the point.

SI 21: K: I think you'd have to have some sort of cut off mechanism when the students were exerting too much force.

SI 22: K: Because some animals, particularly horses when you rectal then can get quite sore. SO some way of when the force is too much... a kind of cut off point.

SI 23: K: And the same sort of thing with duration. There's a tendency with some students, we take them out to the farm, and they can't find the ovaries, and they spend hours and hours and hours trying to find ovaries, and they don't really realise that they've maybe had they're hand in the cow for 10 or 15 minutes. Obviously that's too long, so if there's some kind of time limit you could build into it so that after that you could say if you've not found them it doesn't matter. Keep practicing until you can do.

SI 24: K: As well as actually doing what you have to do, and find out what you have to find out, you've also got to be wary that you're not going to hurt the cow, or if you've got an early pregnancy your going to palpate, there's obviously a risk of abortion, so having some sort of cut off like duration or force.

SI 25: K: I think as well if you could find some way of mimicking a peristalsis you get in a normal cow. So when you were doing it, if you remember the kind of gut movement you got on your arm. If there was some way of mimicking that. Because if the cow was sort of clamped down on kind of rectum around your hand because you're hand's squashed down like this trying to do anything.

SI 26: K: So if there was some way you could adjust the program (to provide variations)...

SI 27: K: as well as (adjusting) what's on the ovary.

SI 28: K: I think it would be use in 1st and 2nd year with the anatomy classes to do it then. But it would also be very useful in 4th and final year. When they do they're clinical stuff. Particularly in final year.

SI 29: K: Yes (There would be time on the course)

Category OP - Ovary Examination Procedure

OP 1: M: Locating the ovaries is important

OP 2: M: Distinguishing between objects (eg intestine, uterus, ovary) by size and consistency of the object.

OP 3: M: Once found, no core skills - just palpate.

OP 4: M: Examinations are mostly ultrasound, although some palpation skills are still needed. Features on a horse ovary are often partially submerged, which means that ultrasound is better at detecting these features

OP 5: M: Cow ovaries are more often palpated.

OP 6: L: faculty does have a code of practice about how many times a rectum can be used a week and that applies to cows and things.

OP 7: L: likeliest thing you would be doing is sequential monitoring the physical characteristics of ovaries to predict when the mare is near to ovulation.

OP 8: L: in the ideal world, you really only want to serve the mare once and that is essentially because every time the stallion enters the mare, he contaminates... the mare is very prone to endemitritous.

OP 9: L: to make sure there's a good chance of pregnancy, you want it to be when she's in ovulation.

OP 10: L: The old fashion way of doing it was that the mare as soon as she was in standing esters would be mated every second day until she was no longer keen on the idea.

OP 11: L: The important thing vets need to know is to be the person that says this mare is going to ovulate within the next 24hours, therefore mate her today.

OP 12: L: So the purpose of predicting ovulation, and that's probably the most important that would have to know.

OP 13: L: now days, that's nearly done exclusively by ultrasound.

OP 14: L: Its actually a relatively small number of people ever will do it to some extent.

OP 15: L: many of the students are the people with the biggest ethical concerns.

OP 16: L: the skill is being able to identify ovaries, identify structures: ovaries and uterus. Appreciate uterine edima, uterine distension. That's probably it.

OP 17: L: likeliest thing that someone would confuse with the things is a ball of faeces.

OP 18: L: Well its (early pregnancy diagnosis) essentially done by ultrasound and the reason its maybe difficult is that there are conditions of the uterine wall that could mimic pregnancy in particular uterine cysts. Because all and early pregnancy looks lie is a kind of Safeway's tomato sized fluid filled lump, so it looks like a black circle about that size.

OP 19: L: The other thing is that I think it's a psychological thing rather than a physical thing. If you say yes it's pregnant of no it isn't, it affects the immediate veterinary management. So the repercussions if you get it wrong are probably greater. It's more difficult to be confident about it. Probably a psychological thing rather than a physical thing. I don't know. I've never tested it. It's just my belief that it's probably a little more difficult (than ovary palpation).

OP 20: L: Well the clever thing you would do is map its uterus before you started breeding so you would know if it happens to have these cysts so that when you now find them three weeks after it had been inseminated, you would know that they were there all along, and they were not the pregnancy. The other thing is you would repeat the ultrasound at an interval of several days, and the ultrasound will progress, the cysts won't.

OP 21: L: it probably feels like apples floating in water and which ones the ovary.

OP 22: L: everyone's grappling around not sure whether they've got a ball of faeces or an ovary.

OP 23: L: What could be confused with pregnancy… a fluid filled bladder an infection within the uterus might feel like it.

OP 24: L: Part of that (infertility diagnosis) would be ovarian assessment, and you could tie it in with that, but there would be other aspects of the investigation such as say uterine biopsy that you know… so it would be a sequential list of questions that people should ask themselves.

OP 25: L: One of the most frustrating things about trying to teach someone to do rectal palpation in general… ovaries is an example, but in any other aspect of rectal examinations like intestines and so on. By definition, they put their hand in, and your trying to talk them through it. And when you say, can you feel this, and half the time they just say yes because they're terrified to say no to me apparently… so they just sort of say it, and you don't know what they're feeling. So you don't know whether they actually got it right or wrong, so you never do know for yourself.

OP 26: L: you just don't know if they are getting it or not.

OP 27: K: Well I guess you'd need to know the actual anatomy. Where things are relative to the ovary. A lot of the problem is actually trying to find the things...

OP 28: K: So its trying to find the ovaries relative to where the uterus is... because your kind of main landmark would be the cervix. You can find the cervix pretty easily in most cases and then you try and work things out from there. So follow it forward to the end of the uterus, and then try and work out where your ovaries are in relation to the uterus. So its kind of knowing where things should be so you can find them I guess.... so manual dexterity I guess would be a core skill.

OP 29: K: It would just be just actually touching it... it would be the actual size of it and kind of consistency of it, and then the sort of surface struct... the surface texture almost.

OP 30: K: I mean is it a kind of smooth bean shaped thing, or has it got something sticking off of it, and if it has, work out what it is.

OP 31: K: follicles are always kind of smooth as opposed to CLs which are very fleshy feel.

OP 32: K: Finding the ovaries (is the hardest thing to teach the students).

OP 33: K: Relatively quickly I suppose. But its just the kind of practice aspect... you know, the more you do the... It's very difficult when the students got they're hand in the cow to try and then... obviously you don't know yourself where it is, so your trying to explain where its most likely to be.

OP 34: K: Its so the students are happy that they can palpate the uterus, and then work out from that where

OP 35: K: Like a heifer that's never had any calves, is going to have a really tiny tract, and really tiny ovaries that are really small. The whole thing fits in your hand. Where as if you've got a suckler cow that's maybe had 10 calves, her tract is going to be bigger anyway, and its a much bigger cow and you've got to stick you're hand in a lot further and grope around a lot more to find it.

OP 36: K: I think it is just really trying to emphasise that it really is like riding a bike. its just really practicing and practicing and practicing and practicing, and doing hundreds and hundreds of cows before you can feel happy enough in your skill of doing it.

OP 37: M: Corpus luteum – Thick walled, ridged

OP 38: M: Follicle – Thin walled, soft

OP 39: M: Ovary – Hard fibrous

Category O - Other Information

O 1: M: You're maybe ignoring other palpation areas. Colic is a big problem in horses, urinary tract palpation. Palpation for a vein is another useful skill.

O 2: M: Needle simulation would not be as useful. A student can practice on an anaesthetised animal. A needle simulation providing different skin types might be useful. For example Pig – rough, cat – soft, horse – soft, cow – tough.

O 3: M: A MIS simulator might also be useful.

O 4: L: it would be relatively easy to adapt in order to teach people other things, and nothing to do with reproductive tracts, in particular diagnosing different types of colic. You could do that because overall the physical methodology and the fact that you've shown that you can use it to appreciate size and textures and that sort of thing... softness, turgidity, that as a proof of concept is good.

O 5: L: That is not to say that access to what you do is not... in fact it would definitely be a very meaningful thing to do in the context of a lab that supported or was complementary to that group of lectures.

O 6: L: if 80% of vets end up in small animal practices what is the point of teaching them those skills.

O 7: L: It's a very small number of people who are actually doing this as a living.

O 8: L: Probably, I would think you could say that 80% of all mare fertility reproductive work is probably done by 100 people in the whole country… well maybe 200.

O 9: L: When Stuart was a student and I, there was a somewhat eccentric or idiosyncratic at least woman who use to teach us reproduction in all species, and she's retired now and she use to… don't put this in any report… she use to particularly in cattle… she's really enthusiastic she tried to get students to feel the right things. And eventually - she couldn't believe they were so incompetent - that she would grab their hand with her hand put it in the cow, and drag you literally to the spot.

O 10: L: I don't know if you know what colic means but they're intestinal disorders that cause abdominal pains. It always forever be – no matter what new technologies come along – it will always be an essential part of a colic examination. And that is a very common scenario that anyone who ends up doing any horse practice will have to do, and actually because it's a naturally occurring disease no matter supposing you have 100 pony herd your not going to get many cases of colic.

O 11: L: So what they're trying to do (during colic examination) is workout has it got distended loops of intestine, and are they small intestinal loops… are they large intestines… are they distended with gas…are they distended with fluid… are they distended with ingesta. That is exactly the questions they go through. It's a simple iterative process.

O 12: L: you could use that (colic sim) again to produce really useful teaching aids. And that would be really useful, and it would be massively fundable.

O 13: L: The other aspect would be all to do with artificial insemination. Which isn't totally unrelated to this subject, because the key point about all this importing frozen semen from Canada at huge expense – well chilled not usually frozen because in cows, its all frozen, and it lasts for year, and you can fly it round the world and all that sort of stuff quite literally. In horses, the synchronisation of the whole thing becomes massively important – in other words, when's it going to ovulate.

O 14: L: (Antony and Rob – Weipers clinical scholars) felt that a whole hand in that involved your whole hand, shoulder movements… they felt and they may be wrong of course… they felt that that would make it more realistic. That depends whether you feel realistic is important because in a way its kind of can you tell the difference between small intestine distension and large intestine… can you tell the difference between gas distension and solid distension because they are fundamental questions in to making the diagnosis.

O 15: L: but small intestinal distension inside a horse feels for all the world as if someone's got a whole load of… the thickness of your wrist… they feel like distended bicycle tubes. You feel as if you're running your hand over a whole load of bicycle tubes.

O 16: L: Because in the solid distended intestine there would be sort of less give, and gassy would be sort of an amorphous thing out of Doctor Who. SO there is a different texture and a different level of resistance.

O 17: L: (A colic sim) would be really useful for students, it would be useful for vets who are doing professional update courses and that sort of thing.

O 18: L: Give texture and shape. I'm presuming shape is possible.

O 19: L: Well what we have just acquired at not inconsiderable expense (for colic training) – Antony maybe mentioned we were getting a physical phantom horse that was suppose to… which is a fibreglass back end of a horse with several sets of intestines that you can stick inside it in different constructions.

O 20: L: Something that might be, and I don't know if its happened yet there was a guy, and its not entirely representative. There was a guy who had a project going on in the States. As far as I know, he was using MRI to kind of create 3D pictures of a horse but because of in MRI scanner you can only get relatively small structures into the MRI scanner, he was actually putting a whole horse in, but they were miniature horses. I mean there are various breeds of horse which are literally that size. It wasn't like a proper big horse.

O 21: L: when you're trying to get them to be able to find a heart murmur. A horse's heart is quite a difficult thing to osciltate. I.e. using a stethoscope. So you kind of say put your stethoscope on here and then say can you here it, and they say yes. And they can't here it at all. In other words, you can never check if they're actually doing it. It's a slightly different thing although in a way its not actually entirely different because their ability to hear it or not is to be able to recognise the noise. Actually there are some crummy recordings of the noise. When you here the whoosh, that's the murmur and there are some people who would be good a going lub-whoosh-dub lub-whoosh-dub, or lub dub whoosh depending on where the murmur comes. But being able to detect it a lot of the time is being able to put their stethoscope on the right part of the horse's heart. The horse's heart is a big thing and it's elbow is in the way and things like that you know where your heart.

O 22: L: Funnily enough, and I'm thinking of third year nightmare classes that we have. The basic physical examination of a horse. And I use to have these crazy objectives on these things, and now if they're able to count the pulse, I consider that to be a major triumph.

O 23: L: It's (horses pulse) quite slow so that puts them off. You literally put their finger on it and say can you feel it now, and they say no. You want to say… well leave the course. You think how can't they feel it. It's under their fingers.

O 24: L: quite a lot of things that are physical feeling like emphysema or pitting edima or… these are things that you show to third year students and say if you press it, it leaves a big pitting…

O 25: L: I mean I remember going to Michigan state vet school, and every single seat in the place has got a PC at it. And the lecturer says Andrew you show what you consider to be the aorta, and you have to show him the aorta….. if you took it a stage further with physical things, or like precordial thrills or fremitus.

O 26: L: . Emphysema is gas under the skin, so it feels like a bubble pack. Ideally, every student should have felt emphysema.

O 27: L: But a lot of the traditional palpation methods are dying out. Which is a shame really but they're… some of them at least that are less important, but some of them will always be important.

O 28: L: Detecting these in physical abnormalities in horses with acute abdominal crises and colic, well the thing that would be good for that would be A) it would be a really really really useful thing to produce, and B) it would be relatively easy to get funds for because there's a big welfare issue. Horse in agony.

O 29: L: your only looking at maybe 5 different kind of overall scenarios (in colic exams). I mean in the ideal world, you have the whole body organs, there, and you have the bits that are normal beside the bits that are not normal. And you'd just put your hand in and say that is the large intestine and it's filled with ingesta. That would be fantastic.

A.4 Differences in Categorisation of Statements

Quote	External Analysis	Initial Analysis
Those are very specific	OP	NA
maybe the mares straining on your arm so you can't actually… that's a slightly different thing	OP	SI
probably the likeliest thing that someone would confuse with the thing is a ball of faeces	SI	OP
it probably feels like apples floating in water	SI	OP
And everyone's grappling around not sure whether they've got a ball of faeces or the ovary	SI	OP
So that would certainly be true	SI	NA
Again, that depends on whether you're trying to use it as a teaching aid.	SI	NA
To be honest, its (post-grad CD) commercially produced and purchased, and its not of sensational quality	CD	CI
The other aspect would be all to do with artificial insemination	CD	NA
We should be doing much more with AI	CD	NA
and numbers do make these things much much more difficult to be able to give them all experience	CI	CD
This faculty does have a code of conduct about how many times a rectum can be used a week and that applies to cows and things	CI	OP
We're not going to try to train out an out	CI	SI
By core skills… the skill is being able to identify ovaries	CI	OP
identify structures: ovaries and uterus	CI	OP
Appreciate uterine edima, uterine distension	CI	OP
But currently, all the students are expected to do exactly the same course, although that's going to change	O	CI
Particularly again when many of the students are the people with the biggest ethical concerns	O	OP
It certainly becomes less fair when you have more students and fewer and fewer cows	O	CD
That is, a guy who makes his living doing (teaching) equine reproduction work and nothing else	O	CI

Obviously you don't palpate sheep ovaries, but you know the kind of theory	O	CI
Finding the ovaries (is the hardest to teach)	CI	OP
But its just the kind of practice aspect… you know the more you do the…	CI	OP
Obviously you don't know where it (the ovary) is, so you're trying to explain where its most likely to be	CI	OP
It's so the students are happy that they can palpate the uterus… and then work out from that where	CI	OP

Table 10. Difference in categorisation of statements betweens the reviewers

A.5 Activity Charts for the Current and Proposed Course

Activity Chart A – The Current Teaching Methods

1. H-H Teacher Delivers a lecture on anatomy and physiology
   H-C The CLIVE system provides graphical and text resources
   O Student learns anatomy & physiology through text book and course notes

2. H-H Student asks a question during a lecture or lab
   H-C
   O

3. H-H Teacher in a lab or lecture responds to a students question
   H-C
   O Book suggests further reading where a more information can be gathered

4. H-H Student discusses with teacher during a lab or asks a question in a lecture
   H-C
   O

5. H-H Teacher presents students with cow tracts in a lab to examine
   H-H Teacher sets student the task to examine the cow reproductive system in-vivo
   H-H Teacher sets student the task to examine the horse reproductive system in-vivo
   O

6. H-H Student attempts the practical and describes what they're feeling to the teacher
   H-C
   O Student describes the procedure in a notebook

7. H-H Teacher attempts to guide student use his/her description of what he/she is feeling
   H-C
   O

8. H-H Student modifies his/her actions depending on the advice given by the teacher
   H-C
   O

9. H-H Student re-evaluates his/her knowledge of the anatomy & physiology with respect to what was experienced in the practical
   H-C
   O

10. H-H Student reconsiders theory to adjust his/her actions at the next practical opportunity
    H-C
    O

11. H-H
    H-C
    O

12. H-H
    H-C
    O

Activity Chart B – The Proposed Training with the Inclusion of a Simulator

1. H-H Teacher Delivers a lecture on anatomy and physiology
   H-C The CLIVE system provides graphical and text resources
   O Student learns anatomy & physiology through text book and course notes

2. H-H Student asks a question during a lecture or lab
   H-C
   O

3. H-H Teacher in a lab or lecture responds to a students question
   H-C
   O Book suggests further reading where a more information can be gathered

4. H-H Student discusses with teacher during a lab or asks a question in a lecture
   H-C
   O

5. H-H Teacher presents students with cow tracts in a lab to examine
   H-H Teacher sets student the task to examine the cow reproductive system in-vivo
   H-H Teacher sets student the task to examine the horse reproductive system in-vivo
   H-C Simulator teaching tool sets student the task to explore the anatomy and physiology of the reproductive system for different features
   O

6. H-H Student attempts the practical and describes what they're feeling to the teacher
   H-C Student attempts to complete the simulator task with the results being entered into the computer either through the interactions or using a mouse and dialog box
   O Student describes the procedure in a notebook

7. H-H Teacher attempts to guide student use his/her description of what he/she is feeling
   H-C Training tool returns performance feedback to the student during the task using current position and force information, and after the exploration using the data entered.
   O

8. H-H Student modifies his/her actions depending on the advice given by the teacher
   H-C **Student modifies his/her actions using the feedback provided by the computer, and attempts to complete the task again**
   O

9. H-H Student re-evaluates his/her knowledge of the anatomy & physiology with respect to what was experienced in the practical

H-C Student revises his/her knowledge of the anatomy and physiology of the equine reproductive system based on the practical simulator task.

O

10. H-H Student reconsiders theory to adjust his/her actions at the next practical opportunity

H-C Student reconsiders the anatomy and Physiology, and adapts the method of performing the simulator task based on this.

O

11. H-H

H-C The simulator provides a different complexity of environment depending on the performance of the student. For example, if the performed the task successfully, introduce more structures into the environment, or use a wider selection of cases.

O

12. H-H

H-C Simulator is altered based on the performance history of students to reflect the areas that most students are having difficulty with.

O

Appendix B: Experiment Sheets and Raw Data from the Experiment Described in Section 5.2

B.1 Introduction

This Appendix contains the data and experiment sheets from the experiment described in Section 5.2. Appendix B.2 shows the answer sheet used during the experiment. Appendix B.3 Shows the NASA TLX sheet used to gather workload data during the experiment. Appendix B.4 contains the raw data collected during the experiment.

B.2 Answer Sheet

Follicle Position Size(cm)
Ovary Side Height

1

1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) 1 2 3 4 5 (High)

2

1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) 1 2 3 4 5 (High)

3

1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) 1 2 3 4 5 (High)

4

1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) 1 2 3 4 5 (High)

Follicle Position Size(cm)
Ovary Side Height

5

1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) 1 2 3 4 5 (High)

6

1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) 1 2 3 4 5 (High)

7

1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) 1 2 3 4 5 (High)

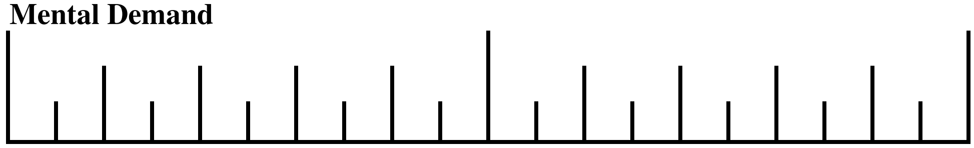
8

1	left/right	front/back	top/bottom	
2	left/right	front/back	top/bottom	
3	left/right	front/back	top/bottom	
4	left/right	front/back	top/bottom	
5	left/right	front/back	top/bottom	

Confidence - (Low) 1 2 3 4 5 (High)

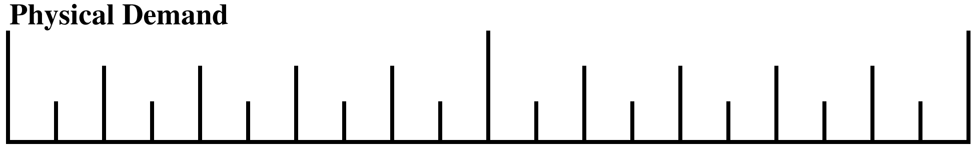
B.3 NASA TLX Scale (with added Confidence and Fatigue scales)

Mental Demand



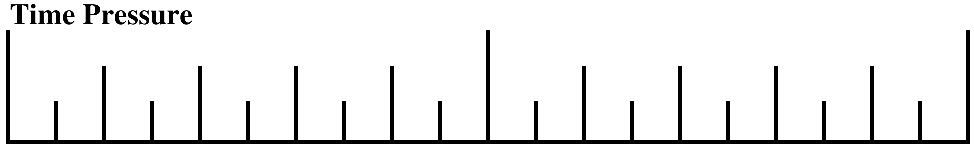
Low High

Physical Demand



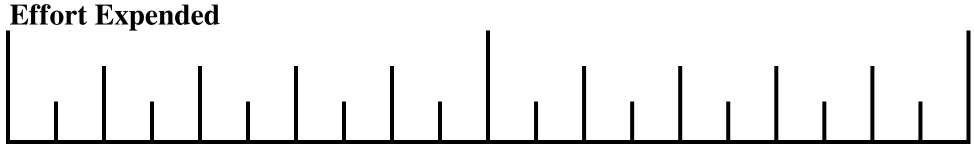
Low High

Time Pressure



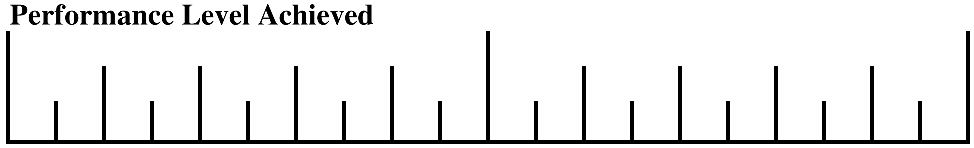
Low High

Effort Expended



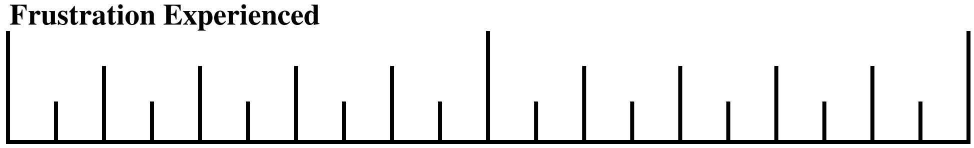
Low High

Performance Level Achieved



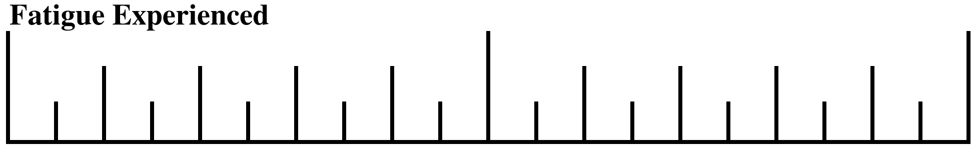
Low High

Frustration Experienced



Low High

Fatigue Experienced



Low High

B.4 Raw Data

Participant	Follicles Placed Correctly	Follicles Placed & Sized Correctly
s1	15	5
s2	16	14
s3	21	10
s4	18	6
s5	13	6
s6	11	4
s7	8	5
s8	19	13
s9	16	8
s10	19	9
v1	8	2
v2	5	1
v3	14	6
v4	21	13
v5	17	11
v6	13	7
v7	17	12

Table 11. Performance data for placing and sizing follicles is shown. The participants are labelled as 'v' veterinarian and 's' for student.

Case	s1	s2	s3	s4	S5	s6	s7	s8	s9	s10
1	199.2	300.0	157.5	154.4	175.1	148.9	300.0	300.0	236.0	265.9
2	199.6	300.0	128.2	101.4	300.0	188.7	300.0	300.0	295.7	300.0
3	142.5	258.8	194.1	105.7	300.0	153.7	300.0	300.0	204.9	300.0
4	160.0	300.0	166.2	140.0	300.0	166.2	300.0	300.0	284.4	300.0
5	237.7	300.0	174.6	110.7	300.0	117.2	300.0	227.0	223.2	200.9
6	116.9	300.0	151.2	300.0	278.7	171.3	197.1	300.0	240.3	300.0
7	148.9	300.0	173.8	160.4	245.9	178.9	300.0	290.1	178.3	300.0
8	149.5	300.0	297.3	100.3	285.7	201.5	300.0	300.0	300.0	238.8

Case	v1	v2	v3	v4	v5	v6	v7
1	259.1	300.0	246.1	183.0	215.0	241.4	140.3
2	238.4	300.0	281.4	228.2	298.9	229.6	218.1
3	300.0	300.0	265.5	300.0	152.1	182.1	300.0
4	300.0	300.0	290.0	300.0	300.0	184.1	192.9
5	300.0	300.0	216.7	300.0	205.1	177.3	180.3
6	300.0	253.0	255.1	158.3	189.8	174.9	230.1
7	171.7	300.0	300.0	216.7	234.8	177.3	300.0
8	228.8	300.0	279.4	181.1	300.0	176.3	248.5

Table 12. Timing data for each examination is shown. The participants are labelled as 'v' veterinarian and 's' for student.

	MD	PD	TP	EE	PA	FrE	FaE	Conf
s1	16	4	5	17	13	13	14	14
s2	14	4	16	16	14	11	14	14
s3	13	9	11	11	11	11	10	11
s4	13	2	5	7	10	12	2	10
s5	16	3	13	8	20	15	9	18
s6	17	6	6	9	7	5	12	7
s7	14	12	12	13	8	12	4	5
s8	12	1	16	11	18	20	7	18
s9	15	8	3	18	18	15	14	16
s10	16	6	12	12	10	16	0	12
v1	17	8	10	16	14	12	12	15
v2	15	7	11	12	10	15	17	11
v3	17	2	10	14	19	8	19	19
v4	14	8	10	12	13	5	12	16
v5	2	6	13	6	13	7	9	12
v6	16	0	0	10	8	2	7	14
v7	13	6	8	12	14	13	15	13

Table 13. Workload data for each participant is shown. The participants are labelled as 'v' veterinarian and 's' for student. The categories are Mental Demand (MD), Physical Demand (PD), Time Pressure (TP), Effort Expended (EE), Performance Achieved (PA), Frustration Experienced (FrE), Fatigue Experienced (FaE), and Confidence (Conf).

	s1	s2	s3	s4	s5	S6	s7	s8	s9	s10
F=0	37247	41472	36258	17336	31860	46618	85245	23942	40992	35411
0<F<=0.3	29417	49144	17791	16493	20903	26963	39675	46200	27610	51148
0.3<F<=0.6	29783	67430	32546	16865	40562	24693	49074	84203	28729	73442
0.6<F<=0.9	25016	50946	31548	24247	56755	17184	36477	58722	31152	45104
0.9<F<=1.2	10619	22007	16709	23760	45636	10232	14494	16764	30313	13685
1.2<F<=1.5	2804	6062	6274	13204	19416	4469	3898	3060	20325	2268
1.5<F<=1.8	442	1548	2144	4359	3925	1715	940	735	10742	343
1.8<F<=2.1	38	317	576	853	628	518	135	137	4160	25
2.1<F<=2.4	23	56	244	190	65	118	30	36	1502	4
2.4<F<=2.7	25	15	95	29	0	46	0	4	491	0
2.7<F<3.0	14	9	44	14	0	28	0	2	211	0
F>=3.0	0	0	37	0	0	37	0	0	138	0

	v1	v2	v3	v4	v5	v6	v7
F=0	32095	45405	37661	47260	31284	37955	44823
0<F<=0.3	50360	38957	41056	18714	35424	24439	62320
0.3<F<=0.6	91307	69640	83592	25125	60286	23270	57540
0.6<F<=0.9	33002	56152	41979	32154	42456	21100	15081
0.9<F<=1.2	2932	21305	7958	28436	18024	19452	1857
1.2<F<=1.5	172	4808	974	17949	3808	14926	121
1.5<F<=1.8	29	841	137	9655	639	7942	21
1.8<F<=2.1	16	337	26	4491	75	3183	18

2.1<F<=2.4	0	85	1	1913	5	1197	0
2.4<F<=2.7	0	21	1	704	7	485	0
2.7<F<3.0	0	1	0	333	1	211	0
F>=3.0	0	0	0	184	16	127	0

Table 14. The number of servo loop cycles spent by each participant in each force range is shown. The participants are labelled as 'v' veterinarian and 's' for student. Force ranges are indicated in Newtons.

	s1	s2	S3	s4	s5	s6	s7	s8	s9	s10
No Contact	37247	41472	36258	17336	31860	46618	85245	23942	40992	35411
Rectal Wall	10302	7034	7956	6098	8003	5920	31562	9164	20307	8560
Left Ovary	30844	53322	43279	25459	72703	23172	39910	66091	44697	42014
Right Ovary	32605	64963	31617	23088	58103	29246	39512	64166	32809	65142
Follicle	24438	72223	25164	45377	49089	27673	33747	70450	57568	70311

	v1	v2	v3	v4	v5	v6	v7
No Contact	32095	45405	37661	47260	31284	37955	44823
Rectal Wall	9863	11104	5417	9842	7928	13362	20044
Left Ovary	66710	59428	43295	43747	50176	28421	36921
Right Ovary	51450	56166	39140	43438	41475	35546	38488
Follicle	49803	65457	87880	42639	61170	39011	41513

Table 15. The number of servo loop cycles spent by each participant in contact with different objects is shown. The participants are labelled as 'v' veterinarian and 's' for student.

	1	2	3	4	5	6	7	8
s1	1.92	1.77	1.68	2.48	1.68	1.94	1.63	2.81
s2	1.78	2.41	2.07	1.92	2.35	2.81	1.92	2.13
s3	1.98	1.83	2.46	3.27	2.59	2.49	1.99	3.3
s4	2.33	2.41	1.95	2.95	2.29	2.71	2.53	1.85
s5	2.28	2.04	2.20	2.09	1.78	1.97	2.14	2.03
s6	2.06	2.59	3.73	2.32	2.23	1.86	2.32	2.15
s7	2.35	1.63	1.69	1.96	1.68	1.90	1.61	1.56
s8	1.66	1.68	2.31	2.52	2.34	1.78	1.62	2.77
s9	3.23	4.00	2.64	2.72	3.36	3.17	2.49	3.23
s10	2.17	1.36	1.96	1.40	1.81	1.51	1.64	1.82
V1	1.40	1.33	1.37	2.02	1.28	1.33	1.82	1.12
V2	2.20	2.40	2.55	2.82	2.09	2.44	2.66	2.63
V3	1.32	2.58	1.51	1.95	2.02	1.67	1.43	1.67
V4	2.66	2.74	3.33	3.66	3.06	3.25	2.96	2.96
V5	1.83	3.79	1.96	1.56	4.61	1.48	2.08	1.85
V6	3.79	3.41	2.95	2.74	3.59	2.85	2.76	2.66
V7	1.99	1.57	1.30	1.48	2.02	1.97	1.37	1.31

Table 16. Peak Forces for all participants in each examination is shown. The participants are labelled as 'v' veterinarian and 's' for student. The force indicated is in Newtons.

Appendix C: Experiment Sheet and Raw Data from the Experiment Described in Section 5.3

C.1 Introduction

This appendix contains the raw data gathered during the experiment described in Section 5.3. Experimental sheets were similar to those used in Appendix B and are therefore not described here. Appendix C.2 contains the raw data gathered from the experiment

C.2 Raw Data

Participant	Session 1	Session 2	Session 3	Session 4	Session 5
S1	13	13	17	15	17
S2	5	8	13	15	11
S3	12	17	16	17	16
S4	11	16	17	17	17
S5	14	15	10	13	16
S6	13	15	17	14	16
S7	13	17	17	17	17
S8	17	17	16	17	17

Table 17. This shows the number of correctly placed follicles for all participants over five training sessions. The maximum number of follicles was in each case 17.

Participant	Session 1	Session 2	Session 3	Session 4	Session 5
S1	5	5	8	9	7
S2	1	5	6	7	7
S3	5	10	12	10	8
S4	6	11	9	10	13
S5	3	3	4	6	8
S6	6	3	12	8	8
S7	9	11	13	11	9
S8	10	6	12	9	12

Table 18. This shows the number of correctly placed and sized follicles for all participants over five training sessions. The maximum number of follicles was in each case 17.

Participant	Session 1	Session 2	Session 3	Session 4	Session 5
S1	268.055625	281.45875	281.173	164.131	187.7525
S2	300	218.959625	245.7258	199.0526	158.7619
S3	284.780625	241.978	248.0679	183.2846	137.6543
S4	203.86925	200.50475	182.1723	159.678	215.2445
S5	205.114875	204.865875	144.2601	119.4214	162.574
S6	241.910125	151.532875	194.6488	212.96	181.4659
S7	208.7875	159.969	151.618	156.7705	144.2374
S8	261.994125	169.232125	144.3353	93.78375	92.27775

Table 19. This shows the mean time taken (in seconds) for each examination for all participants over five training sessions. The maximum time allowed in each case was 300s.

Training Session	Participant	MD	PD	TP	EE	PA	FruE	OW	FatE	CL
1	S1	19	0	4	20	16	18	12.83	4	18
1	S2	18	2	11	14	16	13	12.33	4	17
1	S3	20	3	14	15	16	11	13.17	0	14
1	S4	18	0	2	15	10	0	7.50	0	10
1	S5	16	14	12	16	8	19	14.17	15	4
1	S6	12	1	4	16	7	13	8.83	4	7
1	S7	10	1	4	13	11	10	8.17	1	12

1	S8	16	4	8	12	9	9	9.67	8	8
2	S1	19	1	10	17	9	7	10.50	2	10
2	S2	19	6	7	16	15	16	13.17	4	17
2	S3	12	2	2	10	7	4	6.17	0	8
2	S4	4	0	6	2	6	2	3.33	0	6
2	S5	13	13	6	12	8	9	10.17	7	9
2	S6	15	10	6	15	17	13	12.67	4	15
2	S7	11	5	3	9	6	3	6.17	3	7
2	S8	10	7	5	12	10	5	8.17	2	10
3	S1	16	1	3	19	9	6	9.00	3	8
3	S2	10	2	2	10	8	10	7.00	1	9
3	S3	2	0	2	3	6	1	2.33	1	6
3	S4	17	2	8	17	15	14	12.17	3	14
3	S5	11	11	8	11	9	7	9.50	8	10
3	S6	9	2	2	16	14	5	8.00	4	9
3	S7	11	8	3	11	9	8	8.33	3	10
3	S8	7	7	1	8	4	4	5.17	4	6
4	S1	19	1	4	20	9	6	9.83	4	13
4	S2	17	2	6	18	13	12	11.33	5	15
4	S3	11	1	0	6	14	7	6.50	1	14
4	S4	2	0	0	0	1	0	0.50	0	0
4	S5	11	2	2	10	9	7	6.83	3	9
4	S6	9	13	9	11	14	14	11.67	12	12
4	S7	4	4	1	17	5	2	5.50	3	4
4	S8	4	3	0	3	3	1	2.33	1	3
5	S1	4	0	0	3	4	0	1.83	1	4
5	S2	19	1	3	19	17	17	12.67	7	15
5	S3	14	7	5	14	9	3	8.67	5	10
5	S4	10	11	9	11	10	10	10.17	10	9
5	S5	9	5	5	4	10	8	6.83	4	10
5	S6	13	3	0	3	10	1	5.00	0	10
5	S7	12	5	4	11	10	8	8.33	3	10
5	S8	19	1	6	19	10	13	11.33	1	10

Table 20. This shows the workload data gathered for all participants at the end of each of the five training sessions. The categories are Mental Demand (MD), Physical Demand (PD), Time Pressure (TP), Effort Expended (EE), Performance Achieved (PA), Frustration Experienced (FrE), Overall Workload (OW), Fatigue Experienced (FaE), and Confidence Level (CL).

Appendix D: Experiment Sheet and Raw Data from the Experiment Described in Section 5.4

D.1 Introduction

This appendix contains the experiment sheet and data from the experiment described in Section 5.4. Appendix D.2 contains the answer sheet distributed to each participant. Appendix D.3 contains the raw data gathered from the experiment.

D.2 Answer Sheet Distributed to All Participants

1

Follicle	**Ovary**	**Diameter (cm)**	
1	left/right		
2	left/right		
3	left/right		
4	left/right		
5	left/right		

Confidence - (Low) **1 2 3 4 5** (High)

2

Follicle	**Ovary**	**Diameter (cm)**	
1	left/right		
2	left/right		
3	left/right		
4	left/right		
5	left/right		

Confidence - (Low) **1 2 3 4 5** (High)

3

Follicle	**Ovary**	**Diameter (cm)**	
1	left/right		
2	left/right		
3	left/right		
4	left/right		
5	left/right		

Confidence - (Low) **1 2 3 4 5** (High)

4

Follicle	**Ovary**	**Diameter (cm)**	
1	left/right		
2	left/right		
3	left/right		
4	left/right		
5	left/right		

Confidence - (Low) **1 2 3 4 5** (High)

5

Follicle	**Ovary**	**Diameter (cm)**	
1	left/right		
2	left/right		
3	left/right		
4	left/right		
5	left/right		

Confidence - Low **1 2 3 4 5** High

6

Follicle	**Ovary**	**Diameter (cm)**	
1	left/right		
2	left/right		
3	left/right		
4	left/right		
5	left/right		

Confidence - (Low) **1 2 3 4 5** (High)

7

Follicle	**Ovary**	**Diameter (cm)**	
1	left/right		
2	left/right		
3	left/right		
4	left/right		
5	left/right		

Confidence - (Low) **1 2 3 4 5** (High)

8

Follicle	**Ovary**	**Diameter (cm)**	
1	left/right		
2	left/right		
3	left/right		
4	left/right		
5	left/right		

Confidence - (Low) **1 2 3 4 5** (High)

D.3 Raw Data

Group	Other (Not Follicle)	Extra	Correct Position	Within 0.5	Mean Distance
VR1	7	4	13	4	1.11538
VR2	5	2	13	NA	NA
VR3	10	11	13	7	0.57692
VR4	9	3	14	13	0.28571
VR5	5	5	13	8	0.56154
VR6	2	2	11	9	0.35455
VR7	5	2	13	13	0.23846
VR8	7	4	13	13	0.21154
Trad1	3	2	12	8	0.6
Trad2	3	3	14	13	0.28571
Trad3	6	5	13	12	0.37692
Trad4	5	2	12	8	0.475
Trad5	8	6	13	6	0.88462
Trad6	6	4	12	7	0.4125
Trad7	3	3	7	2	0.75
Trad8	5	4	11	4	0.98227

Table 21. The positioning and sizing follicles gathered during the experiment for each participant is shown. Participants are labelled 'VR' for virtually trained students and 'Trad' for traditionay trained students.

NOTE: VR2 replied with Small, Medium or Large when sizing the follicles rather than providing an actual size. Therefore, these sizing results have been omitted from the table below.

Participant	Mean Time
VR1	267.425
VR2	191.625
VR3	234
VR4	205.625
VR5	204.1
VR6	160.5625
VR7	170.5
VR8	170.45
Trad1	267.2125
Trad2	132.5
Trad3	268.4
Trad4	300
Trad5	240.8
Trad6	209.6125
Trad7	214.525
Trad8	169.925

Table 22. Mean time taken for each examination is shown for all participants. Participants are labelled 'VR' for virtually trained students and 'Trad' for traditionally trained students. The maximum time allowed for an examination was 300s.

Participant	MD	PD	TP	EE	PA	FruE	OW	CL
VR1	19	4	14	19	9	13	13	10
VR2	19	1	6	19	19	15	13.2	18
VR3	14	4	8	10	14	14	10.7	16
VR4	12	3	2	12	17	1	7.8	17
VR5	14	10	6	16	12	12	11.3	12
VR6	10	13	8	12	10	12	10.8	12
VR7	18	14	6	15	14	16	13.8	11
VR8	18	9	4	16	14	7	11.3	14
Trad1	8	2	2	4	16	14	7.7	14
Trad2	14	6	6	13	9	12	10	8
Trad3	14	10	4	16	16	19	13.2	17
Trad4	12	8	10	12	10	14	11	12
Trad5	6	2	9	10	8	16	8.5	8
Trad6	15	4	8	16	10	8	10.2	8
Trad7	7	2	2	10	12	2	5.8	13
Trad8	10	2	4	10	16	4	7.7	14

Table 23. Shown are the workload data gathered for all participants at the end of each of the five training sessions. The categories are Mental Demand (MD), Physical Demand (PD), Time Pressure (TP), Effort Expended (EE), Performance Achieved (PA), Frustration Experienced (FrE), Overall Workload (OW), and Confidence Level (CL).

Appendix E: Experimental Sheets and Data from the Experiment Described in Section 6.3

E.1 Introduction

This appendix contains the result sheet distributed to participants during the final experiment described in this thesis. Appendix E.2 contains the result sheet distributed to each participant for each examination rated. Appendix E.3 shows screen shots of each of the examinations shown in full path mode. Appendix E.4 contains the raw data gathered during the experiment.

E.2 *Results Sheet Distributed To All Participants*

Search Strategy

Has the examiner used an ordered, logic search pattern on the ovaries to look for surface features, or has he/she randomly searched the ovaries?

Random Ordered

Right Ovary Exploration

Has the examiner examined the whole surface of the ovary? There may be sections of the ovary that have not been palpated during the examination that may or may not contain surface features.

None Thorough

Left Ovary Exploration

Has the examiner examined the whole surface of the ovary? There may be sections of the ovary that have not been palpated during the examination that may or may not contain surface features.

None Thorough

Follicle Palpation

Has the examiner searched the ovary but failed to find or identify a surface feature? He/She may also have missed it if he/she touched the feature, but did not recognised it, and therefore did not try to size it.

Follicle	Front Bottom Left	Front Top Right	Back Bottom Right
Successfully Palpated(Y/N)			

Use of too much force

Has the examiner used too much force during the examination (indicated by white)? On how many occasions during the examination did this occur ?

No Times 1 to 3 times 4 to 6 times More than 6 times

Overall Rating

What rating would you give for this examination overall ?

Poor Excellent

Confidence
How confident are you of your answers ?

Not confident | | | | | | | | | | Confident

Examination ____________

E.3 Screenshots of the Examinations in 'Full Path' mode

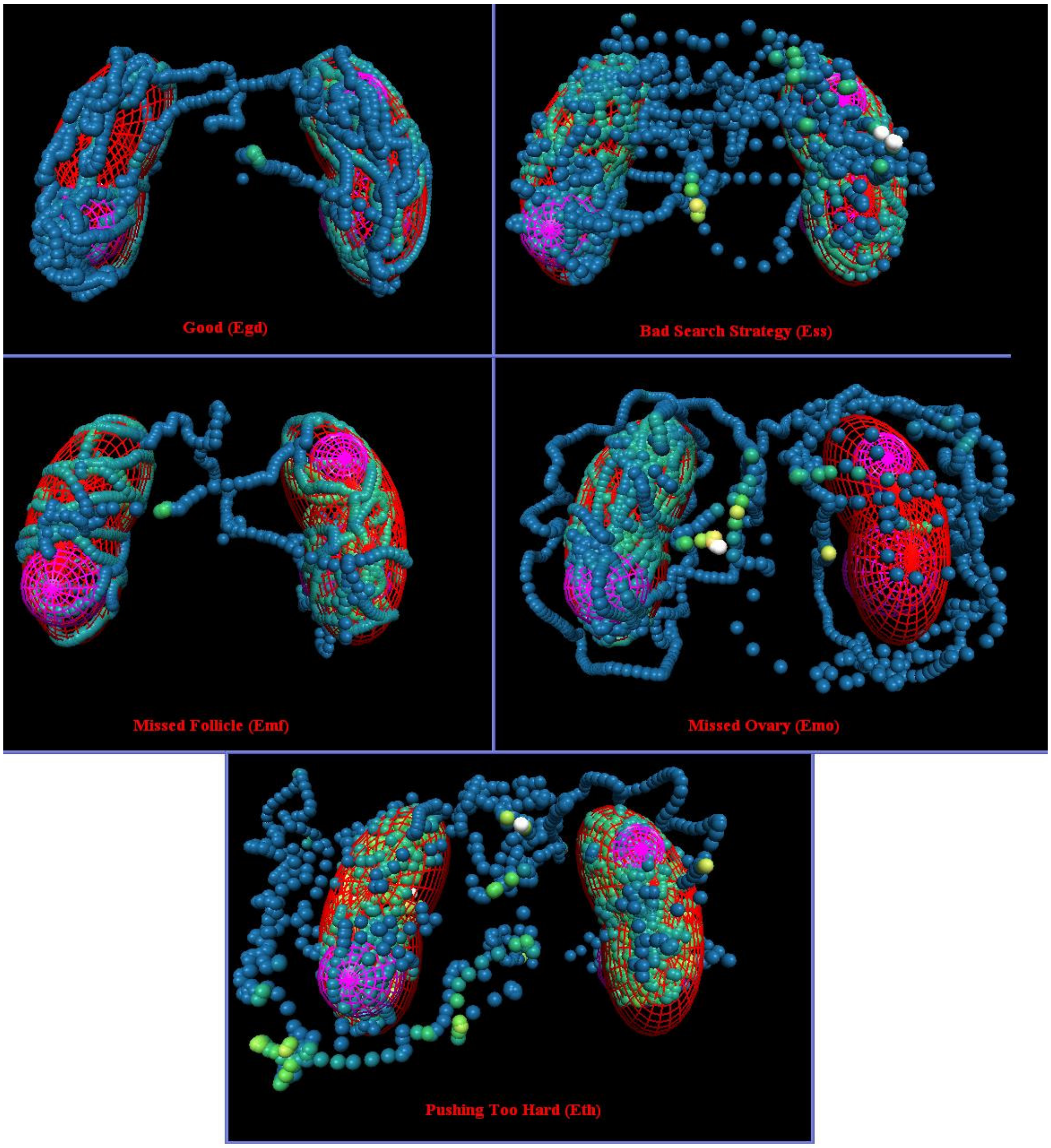


Figure 58. All examinations viewed in 'Full Path' mode.

E.4 Raw Data

Participant	Examination	OR	SS	ROE	LOE	C
1	Egd	9	7	6	9	6
1	Etmf	9	9	5	8	6
1	Emo	2	3	2	10	7
1	Emf	7	6	8	4	5
1	Ess	4	4	3	6	3
2	Egd	9	10	10	10	9
2	Etmf	1	1	7	7	1
2	Emo	5	5	2	9	9
2	Emf	4	8	3	5	8
2	Ess	2	2	3	3	9
3	Egd	9	9	9	9	9
3	Etmf	1	2	2	2	9
3	Emo	4	6	2	10	9
3	Emf	5	8	4	5	9
3	Ess	1	1	2	3	9
4	Egd	7	9	9	3	2
4	Etmf	9	9	9	9	1
4	Emo	1	3	2	9	2
4	Emf	7	9	10	10	7
4	Ess	8	1	9	9	3
5	Egd	4	2	3	3	7
5	Etmf	7	6	7	9	5
5	Emo	2	1	2	5	7
5	Emf	3	2	3	2	8
5	Ess	3	1	4	4	8
6	Egd	8	7	8	8	7
6	Etmf	4	4	6	8	6
6	Emo	3	4	2	6	6
6	Emf	5	5	5	6	7
6	Ess	3	3	5	5	5
7	Egd	9	9	9	7	10
7	Etmf	10	7	6	10	10
7	Emo	5	3	3	10	9
7	Emf	3	9	7	5	8
7	Ess	7	5	6	8	7
8	Egd	8	7	9	9	7
8	Etmf	5	4	3	2	7
8	Emo	3	2	2	2	8
8	Emf	4	3	5	3	7
8	Ess	6	2	6	7	7
9	Egd	10	10	10	10	10
9	Etmf	2	1	3	3	8
9	Emo	3	3	1	6	9
9	Emf	7	8	3	3	7
9	Ess	2	2	2	3	8

10	Egd	8	8	7	4	10
10	Etmf	1	1	1	1	9
10	Emo	1	2	2	4	10
10	Emf	7	9	10	10	5
10	Ess	2	2	1	1	9

Table 24. Shown are the data gathered from all participants for rating each examination. The examinations are described in Chapter 6. The table headings are Overall Rating (OR), Search Strategy (SS), Right Ovary Exploration (ROE), Left Ovary Exploration (LOE), and Confidence (C).

References

[1] Advanced_Medical_Simulations, http://www.medsim.com/

[2] J. G. Arthur, A. D. McCarthy, H. P. Wynn, P. J. Harley, and C. Barber, "Weak at the knees? Arthroscopy Surgery Simulation User Requirements, Capturing the psychological impact of VR Innovation Through risk based design," in Proceedings of Interact 99, Edinburgh UK, 1999, pp. 360-366.

[3] A.-S. Augrelle, M. Penta, O. White, and J.-L. Thonnard, "The Effects of Change in Gravity on the Dynamics of Prehension," in Eurohaptics 2001, Birmingham, UK, 2001, pp. 119-119.

[4] N. J. Avis, N. M. Briggs, D. R. Hose, B. H. Brown, and M. H. Edwards, "Anatomical and Physiological Models for Surgical Simulation," in Proceedings of Medicine Meets Virtual Reality, San Francisco, USA, 1999, pp. 23-29.

[5] V. L. Barker, "Cathsim," in Proceedings of Medicine Meets Virtual Reality, San Francisco, USA, 1999, pp. 36-37.

[6] C. Basdogan, C. Ho, and M. A. Srinivasan, "Simulation of Tissue Cutting and Bleeding for Laparoscopic Surgery Using Auxiliary Surface," in Proceedings of Medicine Meets Virtual Reality, San Francisco, USA, 1999, pp. 38-44.

[7] D. Berg, J. Berkley, S. Weghorst, G. Raugi, G. Turkiyyah, M. Ganter, F. Quintanilla, and P. Oppenheimer, "Issues in Validation of a Dermatologic Simulator," in Medicine Meets Virtual Reality, USA, 2001, pp. 60-65.

[8] D. Berg, G. Raugi, H. Gladstone, J. Berkley, S. Weghorst, M. Ganter, and G. Turkiyyah, "Virtual Reality Simulators for Dermatologic Surgery: Measuring their Validity," *American Society for Dermatologic Surgery*, vol. 27, pp. 370-374, 2001.

[9] J. Berkley, S. Weghorst, H. Gladstone, G. Raugi, D. Berg, and M. Ganter, "Fast Finite Element Modeling for Surgical Simulation," in Proceedings of Medicine Meets Virtual Reality, San Francisco, USA, 1999, pp. 55-61.

[10] L. Bouguila, I. M., and M. Sato, "Scaleable SPIDAR: A Haptic Interface for Human-Scale Virtual Environments," *Lecture Notes in Computing Science*, vol. 2058, pp. 182-193, 2001.

[11] M. Bouzit, G. Popescu, G. Burdea, and R. Boian, "The Rutgers Master II-ND Force Feedback Glove," in IEEE VR 2002 Haptic Symposium, Orland, FL, 2002.

[12] S. A. Brewster, "Visualization tools for blind people using multiple modalities," in Disability and Rehabilitation Technology, 2002, pp. 613-621.

[13] N. M. Briggs and N. J. Avis, "Real Time Soft Tissue Modelling," in Proceedings of the Bizarre Fruit Conference, 1998.

[14] M. Bro-Nielsen, J. L. Tasto, R. Cunningham, and G. L. Merril, "Preop Endoscopic Simulator: A PC-Based Imersive Training System for Bronchoscopy," in Proceedings of Medicine Meets Virtual Reality, San Francisco, USA, 1999, pp. 76-82.

[15] G. Burdea, "Haptic Sensing," in *Force and Touch Feedback for Virtual Reality*: A Wiley-Interscience, 1996, pp. 13 - 40.

[16] G. Burdea, G. Patounakis, V. Popescu, and W. R. E., "Virtual Reality-Based Training for the Diagnosis of Prostate Cancer," *IEEE Transactions on Biomedical Engineering*, vol. 46, pp. 1253-1260, 1999.

[17] J. C. Byers, A. C. Bittner, and S. G. Hill, "Traditional and Raw Task Load Index (TLX) Correlations: Are Paired Comparisons Necessary," *Advances in Industrial Ergonomics and Safety*, vol. 1, pp. 481-485, 1989.

[18] "Rutgers Master II Haptic Interface", CAIP_Rutgers, http://www.caip.rutgers.edu/vrlab/rm2.html

[19] C. Carello and T. M. Turvey, "Rotational Invariants and Dynamic Touch," in *Touch, Representation, and Blindness (Debates in Psychology)*, M. A. Heller, Ed.: Oxford University Press, 2000, pp. 27-66.

[20] A. S. Carlin, H. G. Hoffman, and S. Weghorst, "Virtual reality and tactile augmentation in the treatment of spider phobia: A case study," *Behaviour Research and Therapy*, vol. 35, pp. 153-158, 1997.

[21] H. Coolican, *Research Methods and Statistics in Psychology*, 2 ed: Hodder & Stoughton, 1994.

[22] P. H. Cosman, P. C. Cregan, C. J. Martin, and J. A. Cartmill, "Virtual Reality Simulators: Current Status in Aquisition and Assessment of Surgical Skills," *ANZ Journal of Surgery*, vol. 72, pp. 30-34, 2002.

[23] U. Cugini, M. Bordegoni, C. Rizzi, F. De Angelis, and M. Pratti, "Modelling and Haptic Interaction with non-rigid materials," in Eurographics '99, 1999.

[24] T. Dang, T. M. Annaswamy, and M. A. Srinivasan, "Development and Evaluation of an Epidural Injection Simulator with Force Feedback for Medical Training," in Medicine Meets Virtual Reality, USA, 2001, pp. 97-102.

[25] A. Darzi, S. Smith, and N. Taffinder, "Assessing Operative Skill," *British Medical Journal*, vol. 318, pp. 887-888, 1999.

[26] F. de Angelis, M. Bordegoni, G. Frugoli, and C. Rizzi, "Realistic Haptic Rendering of Flexible Objects," in proceedings of the Second PHANToM Users Group Workshop, 1997, pp. 30-33.

[27] S. De and M. A. Srinivasan, "Rapid Rendering of "Tool-Tissue" Interactions in Surgical Simulations: Thin Walled Membrane Models," in Proceedings of Phantom Users Group, 1998, pp. 23-26.

[28] S. L. Delp, P. Loan, C. Basdogan, and J. M. Rosen, "Surgical Simulation: An Emerging Technology for Training in Emergency Medicine," *Presence*, vol. 6, pp. 147-159, 1997.

[29] J. E. Deutsch, J. Latonio, G. Burdea, and R. Boian, "Rehabilitation of Musculo Skeletal Injuries using the Rutgers Ankle Haptic Interface: 3 case reports," in Eurohaptics 2001, Birmingham, UK, 2001, pp. 11-16.

[30] M. Dinsmore, N. Langrana, G. Burdea, and J. Ladeji, "Virtual Reality Training Simulation of Palpation of Subsurface Tumors," in Proceedings of Virtual Reality Annual International Symposium, Albuquerque, USA, 1997, pp. 54-60.

[31] S. G. Docimo, R. G. Moore, and L. R. Kavoussi, "Telerobotic Surgery is a Clinical Reality: Current Experience with Telementoring in Adults and Children," *Presence*, vol. 6, pp. 173 - 178, 1997.

[32] K. M. Dyce, W. O. Sack, and C. J. G. Wensing, *Textbook of Veterinary Anatomy*: Saunders, 1987.

[33] "Lucky the Virtual Dog", Edinburgh_Virtual_Environment_Centre, http://www.edvec.ed.ac.uk/

[34] P. J. Edwards, A. P. King, D. J. Hawkes, O. Fleig, C. R. Maurer, H. D. L. J., M. R. Fenlon, D. A. de Cunha, R. P. Gaston, S. Chandra, J. Mannss, A. J. Strong, M. J. Gleeson, and T. C. S. Cox, "Stereo Augmented Reality in the Microscope," in Proceedings of Medicine Meets Virtual Reality, San Francisco, USA, 1999, pp. 102 - 108.

[35] M. Enriquez, O. Afonin, B. Yager, and K. Maclean, "A Pneumatic Tactile Alerting System for the Driving Environment," in Workshop on Perceptive User Interfaces, Orlando, 2001, pp.??-??

[36] D. Feygin, M. Keehner, and F. Tendick, "Haptic Guidance: Experimental Evaluation of a Haptic Training Method for a Perceptual Motor Skill," in Haptic Interfaces for Virtual Environment and Teleoperator Systems, Florida, 2002, pp. 40-47.

[37] "The PowerBraille Display", Freedom_Scientific, http://www.freedomscientific.com/fs_products/displays_40.asp

[38] A. G. Gallagher and R. M. Satava, "Virtual Reality as a Metric for the Assessment of Laparoscopic Psychomotor Skills. Learning Curves and Reliability Measures.," *Surgical Endoscopy*, vol. 16, pp. 1746-1752, 2002.

[39] W. W. Gaver, R. B. Smith, and T. O'Shea, "Effective Sounds in Complex Systems: The Arkola Simulation," in Proceedings of CHI 91, 1991, pp. 85-90.

[40] J. J. Gibson, "Observations on Active Touch," *Psychological Review*, vol. 69, pp. 477 - 489, 1962.

[41] E. B. Goldstein, "The Somatic Senses," in *Sensation & Perception*, 4th ed: Brooks/Cole, 1996, pp. 459-499.

[42] P. J. Gorman, J. D. Lieser, W. B. Morray, R. S. Haluck, and T. M. Krummel, "Assessment and Validation of a Force Feedback Virtual Reality Based Surgical Simulator," in Proceedings of Phantom User Group 98, Dedham, Massachusetts, 1998, pp. 27-29.

[43] P. J. Gorman, A. H. Meier, and T. M. Krummel, "Simulation and Virtual Reality on Surgical Education," *Archives of Surgery*, vol. 134, pp. 1203-1207, 1999.

[44] G. A. Higgins, G. L. Merrill, L. J. Hettinger, C. R. Kaufmann, H. R. Champion, and R. M. Satava, "New Simulation Technologies for Surgical Training and Certification: Current Status and Future Projections," *Presence*, vol. 6, pp. 160 - 172, 1997.

[45] R. D. Howe, W. J. Peine, D. A. Kontarinis, and J. S. Son, "Remote Palpation Technology for Surgical Applications," in *Engineering in Medicine and Biology*, vol. 14, 1995, pp. 318-323.

[46] "CyberGrasp", Immersion, http://www.immersion.com/pdf/cybergrasp_datasheet.pdf

[47] "CyberTouch", Immersion, http://www.immersion.com/pdf/cybertouch_datasheet.pdf

[48] "The Laparoscopic Impulse Engine", Immersion, http://www.immersion.com/lapimpulseengine.html

[49] "Laparoscopic Impulse Engine", Immersion, http://www.immersion.com/products/custom/impulseengine.shtml

[50] "MotionMonitor", Innovative_Sports_Training, http://www.innsport.com/

[51] G. Janson and K. Billberger, "The PHANToM Used without Visual Guidance," in proceedings of PHANToM USER RESEARCH SYMPOSUM 99, Heidelberg, Germany, 1999.

[52] G. Jansson, "The Potential Importance of Perceptual Filling-in for Haptic Perception of Virtual Object Form," in Eurohaptics, Birmingham, UK, 2001, pp. 72-75.

[53] K. O. Johnson, R. W. Van Boven, and S. S. Hsiao, "The Peception of Two Points is Not the Spatial Resolution Threshold," in *Touch, Temperature, and Pain in Health and Disease:*

Mechanisms and Assessments, vol. 3, J. Boivie, P. Hansson, and U. Lindblom, Eds. Seattle: IASP Press, 1994, pp. 389-404.

[54] H. Kajimoto, N. Kawakami, T. Maeda, and S. Tachi, "Electro-Tactile Display with Force Feedback," in World Multiconference on Systemics, Cybernetics and Informatics, Orlando, 2001, pp. 95-99.

[55] C. Kauffman, P. Rhee, and D. Burris, "Telepresence Surgery System Enhances Student Training," in Proceedings of Medicine Meet Virtual Reality, San Francisco, USA, 1999, pp. 174-178.

[56] E. Keeve, S. Girod, R. Kikinis, and B. Girod, "Craniofacial Surgery Simulation," in 4th International Conference on Visualization in Biomedical Computing, Hamburg, Germany, 1996, pp. 541-546.

[57] R. L. Klatzky and S. J. Lederman, "Identifying objects from a haptic glance," *Perception & Psychophysics*, vol. 57, pp. 1111-1123, 1995.

[58] R. L. Klatzky, S. J. Lederman, and V. A. Metzger, "Identifying Objects by Touch: An "expert system"," *Perception & Psychophysics*, vol. 37, pp. 299-302, 1985.

[59] L. W. Klein, "Computerized Patient Simulation to Train the Next Generation of Interventional Cardiologists: Can Virtual Reality Take the Place of Real Life," *Catheterization and Cardiovascular Interventions*, vol. 51, pp. 528, 2000.

[60] S. N. Kothari, B. J. Kaplan, E. J. DeMaria, T. J. Broderick, and R. C. Merrell, "Training in Laparoscopic Suturing Skills Using a New Computer-Based Virtual Reality Simulator (MIST-VR) Provides Results Comparable to Those with an Established Pelvic Trainer System," *Journal of Laparoendoscopic & Advanced Surgical Techniques*, vol. 12, pp. 167-173, 2002.

[61] U. Kühnapfel, H. K. Çakmak, and H. Maaß, "3D Modeling for Endoscopic Surgery," in Proceedings of the IEEE Symposium on Simulation, Delft, NL, 1999, pp. 22-32.

[62] R. Kumar, T. M. Goradia, A. C. Barnes, P. Jensen, L. L. Whitcomb, D. Stoianovici, L. M. Auer, and R. H. Taylor, "Performance of Robotic Augmentation in Microsurgery-Scale Motions," *Medical Image Computing and Computer-Assisted Intervention*, vol. LNCS 1679, pp. 1108-1115, 1999.

[63] Y. Kuroda, M. Nakao, S. Hacker, T. Kuroda, H. Oyama, M. Komori, T. Matsuda, and T. Takahashi, "Haptic Force Feedback with an Interaction Model between Multiple Deformable Objects for Surgical Simulation," in Eurohaptics, Edinburgh, UK, 2002, pp. 116-121.

[64] N. Langrana, G. Burdea, J. Ladeji, and M. Dinsmore, "Human Performance Using Virtual Reality Tumor Palpation Simulation," *Computers and Graphics*, vol. 21, pp. 451-458, 1997.

[65] D. Laurillard, *Rethinking University Teaching: a framework for effective use of educational technology*: Routledge, 1993.

[66] S. J. Lederman and R. Klatzky, "Hand Movements: A Window into Haptic Object Recognition," *Cognitive Psychology*, vol. 19, pp. 342-368, 1987.

[67] S. J. Lederman and R. L. Klatzky, "Action for Perception: Manual Exploration and Haptic Processing," in *Hand and Brain: The Neourophysiology and Psychology of Hand Movements*, A. M. Wing, Ed.: Academic Press, 1996, pp. 431-446.

[68] H. Levkowitz, "Perceptual Steps Along Colour Scales," *International Journal of Imaging Systems and Technology*, vol. 7, pp. 97-101, 1996.

[69] M. Liévin and E. Keeve, "Stereoscopic Augmented Reality System for Computer Assisted Surgery," in Computer Assisted Radiology and Surgery, Berlin, Germany, 2001, pp. 27-30.

[70] R. W. Linderman and J. N. Templeman, "Vibrotactile feedback for handling virtual contact in immersive virtual environments," in *Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents and Virtual Reality*: Lawrence Erlbaum Assoc, 2001.

[71] R. Loureiro, F. Amirabdollahian, S. Coote, E. Stokes, and W. Harwin, "Using Haptics Technology to Deliver Motivational Therapies in Stroke Patients: concepts and initial pilot studies," in Eurohaptics, Birmingham, UK., 2001, pp. 1-6.

[72] T. H. Massie and K. Salisbury, "The Phantom Haptic Interface: A Device for Probing Virtual Objects," in Proceedings of the ASME International Mechanical Engineering Congress and Exhibition, Chicago, IL, 1994, pp. 295-302.

[73] A. McCarthy, P. Harley, and R. Smallwood, "Virtual Arthroscopy Training Do the virtual skills developed match the real skills required?," in Proceedings of Medicine Meets Virtual Reality, San Francisco, USA, 1999, pp. 221-227.

[74] M. R. McGee, "Investigating a Multimodal Solution for Improving Force Feedback Generated Textures," in *Computing Science*. Glasgow: University of Glasgow, 2003, pp. 161.

[75] A. M. Meier, C. L. Rawn, and T. M. Krummel, "Virtual Reality: Surgical Application - Challenge for the New Millenium," *American College of Surgeons*, vol. 192, pp. 372-384, 2001.

[76] Mentice_Medical_Simulation, http://www.mentice.com/frameset.html

[77] B. Michell, "Revalidation and Virtual Patients: a Vision of the Future," *(In Practice) Journal of Veterinary Postgraduate Clinical Studies*, vol. 24, pp. 221-223, 2002.

[78] M. Minsky, O. Ming, O. Steele, F. P. Brooks, and M. Behensky, "Feeling and seeing: issues in force display," in Proceedings of the 1990 symposium on Interactive 3D graphics, NY, USA, 1990, pp. 235-241.

[79] M. Montgomery-Masters, "A Laurillardian CAL Design Method - Developed and Implemented," in Proceedings of ED-MEDIA, Freiburg, 1998, pp. 1746-1747.

[80] M. Montgomery-Masters, "Activity-Based CAL Design: A theoretically based design method for CAL materials," in *Computing Science*. Glasgow: University of Glasgow, 2000, pp. 148.

[81] J. N. Moore, T. Melton, W. C. Carter, A. L. Wright, and M. L. Smith, "A New Look at Equine Gastrointestinal Anatomy and Selected Intestinal Displacements," in British Equine Veterinary Association 2002, Glasgow, UK., 2002, pp. 25-26.

[82] F. Moroney, D. W. Biers, and F. T. Eggermeier, "Some Measurement and Methodological Considerations in the Application of Subjective Workload Measurement Techniques," *The International Journal of Aviation Psychology*, vol. 5, pp. 87-106, 1995.

[83] National_Institute_of_Standards_and_Technology, "Virtual Environments for Health Care - A White Paper for the Advanced Technology Program (ATP) National Institute of Standards and Technology," National Institute for Standards and Technology, Gaithersberg, USA October 1995.

[84] "The Visible Human Project", National_Library_of_Medicine, http://www.nlm.nih.gov/research/visible/visible_human.html

[85] V. R. Neufeld and G. R. Norman, *Assessing Clinical Competence*. New York: Springer, 1985.

[86] NovInt_Technologies, http://www.novint.com/

[87] I. Oakley, "Haptic Augmentation of the Cursor: Transforming Virtual Actions into Physical Actions," in *Computing Science*. Glasgow: University of Glasgow, 2003, pp. 283.

[88] I. Oakley, M. R. McGee, S. A. Brewster, and P. Gray, "Putting the Feel in 'Look and Feel'," in proceedings of CHI 2000, The Hague, 2000.

[89] R. O'Toole, R. Playter, W. Blank, N. Cornelius, W. Roberts, and M. Raibert, "A Novel Virtual Reality Surgical Trainer with Force Feedback Surgeon Vs Medical Student performance," in proceedings of Phantom User Group 97, 1997, pp. 73-75.

[90] M. Ottensmeyer and J. K. Salisbury, "Hot and Cold Running VR: adding thermal stimuli to the haptic experience," in PHANToM Users Group, MA, USA, 1997, pp. 39-42.

[91] A. M. Paisley, P. J. Baldwin, and S. Patterson-Brown, "Validity of Surgical Simulation for the Assessment of Operative Skill," *British Journal of Surgery*, vol. 88, pp. 1525-1532, 2001.

[92] F. E. Pollick, C. Chizk, C. Hager-Ross, and M. Hayhoe, "Implicit Accuracy Constraints in Two-Fingered Grasps of Virtual Objects with Haptic Feedback," *Haptic Human Computer Interaction*, vol. LNCS 2058, pp. 98-107, 2001.

[93] A. Radetzky, M. Rudolph, W. Stefanon, S. Starkie, B. Davies, and L. M. Auer, "Simulating Minimally Invasive Neurosurgical Interventions using an Active Manipulator," *Medical Image Computing and Computer-Assisted Intervention*, vol. LNCS 1935, pp. 578-587, 2000.

[94] G. Reinhart, O. Anton, M. Ehrenstrasser, and C. Patron, "Telepresent Microassembly," in PHANToM Users Research Symposium, Zurich, Switzerland, 1999, pp. 9-17.

[95] C. Richard and M. R. Cutkosky, "Contact Force Perception with an Ungrounded Haptic Interface," in ASME IMECE 6th Annual Symposium on Haptic Interfaces, Dallas, 1997, pp. 15-21.

[96] J. M. Rolfe and K. J. Staples, *The Flight Simulator as a Training Device*: Cambridge University Press, 1986.

[97] C. M. S. Royston, M. R. J. Lansdown, and W. A. Brough, "Teaching laparoscopic surgery: the need for guidelines," *British Medical Journal*, vol. 308, pp. 1023-1025, 1994.

[98] M. Sakuma, S. Masamori, T. Harada, J. Sahrhage, and J. Blauert, "System for Japanese Calligraphy Lesson with Force Feedback," in Eighth Annual Symposium on HAPTIC INTERFACES FOR VIRTUAL ENVIRONMENT AND TELEOPERATOR SYSTEMS, 1999.

[99] J. K. Salisbury, "Making Grapics Physicaly Tangible," in *Communications of the ACM*, vol. 42, 1999, pp. 75-81.

[100] E. L. Sallnäs and J. Kjoberg, "Handing Over Objects in a Haptic Collaborative Environment," in Eurohaptics, Edinburgh, UK, 2002, pp. 77-81.

[101] R. M. Satava, "Virtual reality and telepresence for military medicine," *Annals Academy of Medicine Singapore*, vol. 26, pp. 118-120, 1997.

[102] R. M. Satava and S. B. Jones, "Virtual Environments for Medical Training and Education," *Presence*, vol. 6, pp. 139-146, 1997.

[103] SensAble_Technologies, http://www.sensable.com

[104] "PHANToM 1.5 6DOF", SensAble_Technologies, http://www.sensable.com/products/datafiles/phantom_ghost/PHANTOM_GHOST%20Brochure.pdf

[105] K. P. Sherman, J. W. Ward, D. P. M. Wills, and A. M. M. A. Mohsen, "A Portable Virtual Environment Knee Arthroscopy Training System with Objective Scoring," in Proceedings of Medicine Meets Virtual Reality, San Francisco, USA, 1999, pp. 335-336.

[106] K. P. Sherman, J. W. Ward, D. P. M. Wills, V. J. Sherman, and A. M. M. A. Mohsen, "Surgical Trainee Assessment using a VE Knee Arthroscopy Training System (VE-KATS): Experimental Results," in Medicine Meets Virtual Reality, USA, 2001, pp. 465-470.

[107] K. B. Shimoga, "A Survey of Perceptual Feedback Issues in Dexterous Telemanipulation: Part II Touch Feedback," in Proceedings of IEEE International Symposium on Virtual Reality, Seattle, WA, 1993.

[108] C. M. Smith, "Human Factors in Haptic Interfaces," in *Crossroads*, 1999.

[109] S. Smith, A. Wan, N. Taffinder, S. Read, R. Emery, and A. Darzi, "Early Experience with Procedicus VA - The Prosolvia Virtual Reality Arthroscopy Shoulder Trainer," in Proceedings od Medicine Meets Virtual Reality, San Francisco, 1999, pp. 337 - 343.

[110] J. Stalfors, T. Kling-Peterson, M. Rydmark, and T. Westin, "Haptic Palpation of Head and Neck Cancer Patients - Implications for Education and Telemedicine," in Medicine Meets Virtual Reality, USE, 2001.

[111] R. J. Stone, "Haptic Feedback: A Brief History from the Telepresence to Virtual Reality," *Lecture Notes in Computing Science*, vol. 2058, pp. 1-16, 2001.

[112] H. Z. Tan, "Haptic Interfaces," in *Communications of the ACM*, vol. 43, 2000, pp. 40-41.

[113] F. Tendick, M. Downes, T. Goktegin, M. C. Cavusoglu, D. Feygin, X. Wu, R. Eyal, M. Hegarty, and L. W. Way, "A Virtual Environment Testbed for Training Laparoscopic Surgical Skills," *Presence*, vol. 9, pp. 236-255, 2000.

[114] R. Q. Van der Linde, P. Lammertse, E. Frederiksen, and B. Ruiter, "The HapticMaster, a new high-performance haptic interface," in Eurohaptics, Edinburgh (UK), 2002, pp. 1-5.

[115] H. A. H. C. van Veen and J. B. F. van Erp, "Tactile Information Presentation in the Cockpit," *Haptic Human-Computer Interaction*, vol. LNCS 2058, pp. 50-53, 2001.

[116] "VTPlayer", VirTouch, http://www.virtouch.com

[117] S. Wall, "An investigation of temporal and spatial limitations of haptic interfaces," in *Cybernetics Department*. Reading: University of Reading, 2001, pp. 146.

[118] S. Wall and S. A. Brewster, "Assessing Haptic Properites for Data Representation," in CHI 2003, Fort Lauderdale, Florida, 2003, pp. 858-859.

[119] S. Wall and W. Harwin, "Design of a Multiple Contact Point Haptic Interface," in Eurohaptics, Birmingham, UK, 2001, pp. 146-148.

[120] K. Wegner, "Surgical Navigation System and Method Using Audio Feedback," in Proceedings of ICAD'98, Glasgow, Scotland, 1998, pp. 6.2.

[121] Y. Yokokohji, R. L. Hollis, T. Kanade, K. Henmi, and T. Yoshikawa, "Toward Machine Mediated Training of Motor Skills -Skill Transfer from Human to Human via Virtual Environment," in proceedings of RO-MAN, Japan, 1996, pp. 32-37.